CHAPMAN UNIVERSITY
ADMINISTRATION

DANIELE C. STRUPPA
President

NORMA BOUCHARD, PH.D.
Executive Vice President, Provost and Chief Academic Officer

HAROLD W. HEWITT, JR.
Executive Vice President and Chief Operating Officer

MATTHEW PARLOW
Executive Vice President and Chief Advancement Officer

BRIAN K. POWELL, JD
Vice President and Chief Human Resources Officer

COLLETTE CREPELL
Vice President of Campus Planning and Design

HELEN NORRIS
Vice President and Chief Information Officer

JAMIE CEMAN
Vice President of Strategic Marketing and Communications

MARTINA NIESWANDT, PH.D
Vice President for Research

JANNA BERSI, MBA, ED.D.
Vice President of Investments and Administration

JERRY PRICE, PH.D.
Vice President for Student Affairs and Dean of Students

JOSEPH M. SLOWENSKY
Vice Provost for Institutional Effectiveness and Faculty Affairs

MIKE PELLY
Vice President and Dean of Enrollment

REG CHHEN STEWARD, PH.D.
Vice President of Diversity, Equity and Inclusion

AMY ROGAN-MEHTA
Vice President of University Advancement

BOARD OF TRUSTEES

PARKER S. KENNEDY, Chair
JAMES V. MAZZO, Vice Chair
SCOTT CHAPMAN, Secretary
MARILYN ALEXANDER
LISA ARGYROS '07
RAJ S. BHATHAL
GAYE BIRTCHER
KEN BUNT '93
JAMES P. BURRA
ROBERT CASE
AKIN CEYLAN '90
IRVING M. CHASE
JEROME W. CWIERTNIA
ZEINAB DABBH (JD '12)
DALE E. FOWLER '58
EMILE HADDAD
GAVIN S. HERBERT, JR.
MARK HILBERT
WILLIAM K. HOOD
ANDY HOROWITZ
ANDRÉS IRLANDO
MARK CHAPIN JOHNSON '05
CHERYL LENTZ
RICK LOWE
MELINDA MASSON
MARONYA MOULTREIE
MARYBELLE MUSSCO
RICHARD MUTH (MBA '81)
JAMES B. ROSZAK  
THE HONORABLE LORETTA  
SANCHEZ ‘82  
MOHINDAR S. SANDHU  
RONALD M. SIMON  
MAX VALDES  
TIM VANDERHOOK  
THE HONORABLE GADDI H.  
VASQUEZ ’09  
ANNETTE WALKER  
GEORGE WALL  
NELLA WEBSTER-O’GRADY ‘71  
KAREN R. WILKINSON ’69  
JANE FUJISHIGE YADA  
CHARLIE ZHANG

EMERITUS CHAIRS
WYLIE A. AITKEN  
THE HONORABLE GEORGE L.  
ARGYROS ’59  
DOY B. HENLEY  
DONALD E. SODARO

EMERITUS TRUSTEES
DONNA FORD ATTALLAH ’61  
ARLENE R. CRAIG  
DAVID C. HENLEY  
ROGER C. HOBBS  
JOANN LEATHERBY  
TOM MALLOY  
RICHARD R. SCHMID  
R. DAVID THRESHIE

EX-OFFICIO TRUSTEES
REVEREND LT AUDUI A BYNUM ’76  
AARON FLEWELLEN ’04  
RALPHIE GIRON ’14  
REVEREND JAY HARTLEY  
REVEREND DAYNA KINKADE  
PAULA MCCANCE  
SHERI NAZAROFF  
MICHAEL PENN  
REVEREND RICHIE SANCHEZ  
DANIELE C. STRUPPA

BOARD OF ADVISORS
THOMAS D. PHELPS, Chair  
WYLIE A. AITKEN  
ROBERT ALVARADO, JR.  
PHILLIP H. CASE  
JOHN R. EVANS  
WOLFGANG FRISCH ’97 (JD ’00)  
THE HONORABLE ANDREW J.  
GUILFORD  
DOY B. HENLEY  
JANET E. HUMPHREY  
PARKER S. KENNEDY  
THE HONORABLE LAYNE H. MELZER  
DAVID MURPHY

SCHOOL OF LAW ADMINISTRATION
MARISA S. CIANCIARULO  
Interim Dean and Donald P.  
Kennedy Chair in Law, Doy and  
Dee Henley Chair in Law, and  
Professor of Law

JENNY CAREY  
Associate Dean for Academic  
Affairs, Professor of Legal  
Research and Writing

RICHARD E. REDDING  
Associate Dean for Research and  
Faculty Development

MARIO MAINERO  
Associate Dean for Academic  
Achievement and Bar Preparation

JUSTIN CRUZ  
Assistant Dean of Admission and  
Diversity Initiatives

SUSIE PARK  
Assistant Dean for Career Services

KELLY FARANO  
Assistant Dean for Administration
<table>
<thead>
<tr>
<th>Name</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARIRA A. SADEGHI</td>
<td>Sam &amp; Ash Assistant Dean of Academic Achievement</td>
</tr>
<tr>
<td>GEORGE WILLIS</td>
<td>Director of the Tax Law Clinic, Clinical Professor, and Administrator of the Graduate Tax Programs, and Professor of Law</td>
</tr>
<tr>
<td>MARYAM ISLES</td>
<td>Assistant Dean for Academic Services and Registration</td>
</tr>
<tr>
<td>CAMILLE HEENAN</td>
<td>Assistant Dean for Student Affairs</td>
</tr>
<tr>
<td>LAW SCHOOL FACULTY</td>
<td></td>
</tr>
<tr>
<td>Deepa Badrinarayana</td>
<td>Professor of Law</td>
</tr>
<tr>
<td>Michael Bazyler</td>
<td>Professor of Law and 1939 Society Law Scholar in Holocaust and Human Rights Studies</td>
</tr>
<tr>
<td>Tom W. Bell</td>
<td>Professor of Law</td>
</tr>
<tr>
<td>Sean Bigley</td>
<td>Assistant Professor of Law, Visiting Faculty</td>
</tr>
<tr>
<td>Denis Binder</td>
<td>Professor of Law</td>
</tr>
<tr>
<td>Daniel Bogart</td>
<td>Professor of Law and Bolinger Chair in Real Estate, Land Use and Environmental Law</td>
</tr>
<tr>
<td>Dr. Thomas Campbell</td>
<td>Professor of Law, Doy and Dee Henley Distinguished Professor of Jurisprudence, and Former Dean (2011–2016)</td>
</tr>
<tr>
<td>Lan Cao</td>
<td>Professor of Law and Betty Hutton Williams Professor of International Economic Law</td>
</tr>
<tr>
<td>Jenny Carey</td>
<td>Associate Dean of Academic Affairs, Professor of Legal Research and Writing</td>
</tr>
<tr>
<td>Marisa S. Cianciarulo</td>
<td>Interim Dean and Donald P. Kennedy Chair in Law, Doy and Dee Henley Chair in Law, and Professor of Law</td>
</tr>
<tr>
<td>Bobby Dexter</td>
<td>Professor of Law</td>
</tr>
<tr>
<td>Kurt Eggert</td>
<td>Professor of Law, Director of the Alona Cortese Elder Law Center</td>
</tr>
<tr>
<td>Judd Funk</td>
<td>Professor of the Practice of Entertainment Law and Director of the Entertainment On-Ramp Program</td>
</tr>
<tr>
<td>Dr. John A. Hall</td>
<td>Professor of Law and Director of International Law Emphasis Program</td>
</tr>
<tr>
<td>Kathy Heller</td>
<td>Associate Professor of the Practice of Law and Executive Director of the Entertainment Law Emphasis Program</td>
</tr>
<tr>
<td>Ernesto Hernandez</td>
<td>Professor of Law</td>
</tr>
<tr>
<td>Hugh Hewitt</td>
<td>Professor of Law</td>
</tr>
</tbody>
</table>
ARTICLES

A Content Analysis of Criminal Cases Concerning Unattended Children in Vehicles Between 1990 and 2021: Empirically-Based Suggestions for Reform
Jacob W. Forston and Henry F. Fradella ......................... 1

In the Name of the Environment Part III: CEQA, Housing, and the Rule of Law
Jennifer Hernandez ........................................................ 57

The Foreign Tax Credit Redux
Bret Wells ............................................................... 159

TRANSCRIPT

Chapman Law Review Debate: Does Originalism Work?
Kurt Eggert and Lee Strang, moderated by Tom Campbell................................. 237

NOTES

H-1B Visa Lottery: Random Selection Process Prevents Applicants from Achieving the American Dream
Alexandra Amos .......................................................... 263

Challenging Subpar Servitudes
Samantha Kuo............................................................ 297
A Case for Protecting Youth from the Harmful Mental Effects of Social Media
  Kaidyn McClure ............................................................ 325

Physician False Claims Act Liability—The Circuit “Split” That Illustrates the Need for Health Courts
  Nicole Rickerd ................................................................. 361

Stars on the Field, Benchwarmers on the Tax Return: Student-Athletes and the Tax Ramifications of Name, Image, and Likeness Deals
  Haley A. Ritter ................................................................. 401
Editor’s Note

Chapman Law Review is pleased to release the first Issue of Volume Twenty-Six. This Issue comprises scholarship covering a diverse range of subjects across numerous legal domains.

In the first article, Mr. Jacob W. Forston and Dr. Henry F. Fradella present an empirical analysis of children left unattended in vehicles over the course of thirty-one years. Based on their findings and analysis, the authors suggest reforms to improve the social control of the underlying behaviors that have potentially lethal consequences. The second article, written by Ms. Jennifer Hernandez, is the third study in her series, In the Name of the Environment, providing a thorough evaluation of the litigation surrounding the California Environmental Quality Act. Next, Professor Bret Wells examines the amendment to Section 901—regarding the foreign tax credit—in the U.S. Treasury Department’s 2022 final regulations and suggests additional amendments to rectify inconsistencies with historic policy goals of the credit.

This Issue then provides a transcript of a riveting debate between Professor Kurt Eggert and Professor Lee Strang, moderated by Professor Tom Campbell, regarding the topic: Does Originalism Work? This debate was a response to Professor Eggert’s previously published article: Originalism Is Not What It Used to Be.

The remaining pieces in this Issue are notes written by J.D. Candidates currently in their third year at Chapman University Dale E. Fowler School of Law and members of the Chapman Law Review. The first note, written by Ms. Alexandra Amos, addresses the faults in the H-1B Visa Lottery and proposes a merit-based system as a solution for immigrants to achieve the American Dream. In Ms. Samantha Kuo’s note, she argues that golf courses should not be protected by conservation easements and that California’s housing crisis could be improved by building housing on these golf courses. Ms. Kuo suggests a change to the legal

---


landscape on conservation easements to enable the termination of such conservation easements on golf courses. The next note, written by Ms. Kaidyn McClure, assesses possible legal approaches to protect teens from the mental health harms stemming from social media. Then, Ms. Nicole Rickerd’s note addresses the alleged circuit split on physician liability under the False Claims Act, ultimately determining the disagreement is a misunderstanding and not an actual split. Ms. Rickerd insists that there is a need for specialized health courts which would have jurisdiction over all civil federal healthcare cases, including addressing the circuit “split” at issue. I authored the final piece in this Issue, addressing the tax ramifications from the Supreme Court case, *NCAA v. Alston*,3 and other recent laws allowing student-athletes to receive income for their name, image, and likeness.

*Chapman Law Review* expresses profound gratitude towards the faculty and administration who have contributed to the realization of this Journal. Particularly, we are immensely appreciative of our faculty advisor, Professor Celestine McConville, who has been an invaluable asset throughout the Journal creation process, offering guidance and expertise at every stage. Furthermore, we extend our gratitude to Interim Dean of Chapman University Dale E. Fowler School of Law and Professor of Law, Dean Marisa S. Cianciarulo, and our esteemed faculty advisor committee, including Professor Kenneth Stahl, Professor Nancy Schultz, and Professor Carolyn Larmore. We also wish to acknowledge our gratitude to the Research Librarians of the Hugh & Hazel Darling Law Library, whose expertise has been a vital resource for source collection. Furthermore, I would like to acknowledge the invaluable contributions of our Executive Managing Editor, Rachel McMains, and our Executive Production Editor, Sarah McMillin, for their unwavering commitment to the production of an exemplary publication. Finally, I wish to express my utmost gratitude for the opportunity to collaborate with the 2022-2023 *Chapman Law Review* editorial team. Working with you all has been an absolute honor, and I take immense pride in the accomplishments we have achieved this year. I am humbled to have been a part of this incredible team, and I am grateful for your unwavering commitment and diligent efforts.

Haley A. Ritter

*Editor-in-Chief*

---

A Content Analysis of Criminal Cases Concerning Unattended Children in Vehicles Between 1990 and 2021: Empirically-Based Suggestions for Reform

Jacob W. Forston† and Henry F. Fradella‡

Leaving children unattended in vehicles is one of the leading causes of vehicle deaths not associated with a crash. Intended deaths that are purposefully caused via this phenomenon are quite rare. Rather, such fatalities are typically a result of a caregiver either forgetting a child in a vehicle or making a conscious decision to leave the child unattended without realizing the dangers attendant to that decision. Either way, the resultant harm sparks moral outrage in the media and the community. This, in turn, can prompt prosecution of caregivers under circumstances in which their actions may not align with the elements of the crimes with which they are charged.

† This Article was written during the summer of 2022. The statistics reported for children left unattended in vehicles were current as of the dates indicated in “last visited” parentheticals in applicable footnotes. Incidents occurring after August 5, 2022, are not included in the data presented in this Article.

† B.S. and M.S. in criminology and criminal justice, Arizona State University. Mr. Forston’s research interests include legal decision-making, specifically plea decision-making. Mr. Forston wrote an earlier version of this paper during the spring 2022 semester in partial satisfaction of the requirements for Dr. Fradella’s graduate seminar in criminal law and social control.

This Article synthesizes the extant literature on the phenomenon of children left unattended in vehicles (“CLUV”) by examining the dangers associated with the behavior and both the prevalence and conditions under which the phenomenon occurs. The Article then analyzes the laws that some state legislatures enacted to curb the CLUV phenomenon, focusing on criminal legal responses. By conducting an original, mixed-method content analysis of cases from across the United States, the Article presents a typology of how courts adjudicate CLUV cases that include both pediatric hyperthermia fatalities and those in which children survived CLUV incidents. Qualitative analysis reveal three overarching themes in CLUV cases, including those that involve disputes regarding the sufficiency of the evidence (often focusing on mens rea), questions of statutory construction, and challenges to the collateral consequences of convictions. Quantitative analyses demonstrate that across these three themes, the prosecution prevails in CLUV cases by a ratio of more than two to one. This appears to be due, in part, to the fact that in roughly one out of every five cases, the caregiver’s actions were attendant to either being under the influence of alcohol or other drugs at the time of the CLUV incident or participating in other criminal behaviors. Nonetheless, the cases in the research sample had a 32.6% reversal rate for sufficiency of the evidence claims—a rate quadruple that of the national reversal rate for all other crimes challenged on appeal on such grounds. This finding, in turn, suggests that prosecutors should rethink their approaches to CLUV cases. Additionally, legislatures could take steps to clarify the elements of CLUV-related offenses. Toward that end, the Article offers a suggested statute that would address the questions raised in the cases analyzed in this research. Finally, the Article concludes by offering and alternative ways to address children being harmed while unattended in vehicles using both formal and informal social controls.
C. CLUV-Specific Laws ........................................... 12
D. Charging and Prosecuting Decisions in Pediatric
   Hyperthermia Death Cases ............................... 15
   Table 3: Case Outcomes in CLUV Death
   Cases..................................................................... 16

II. METHODS ................................................................. 16
   A. Removal of Irrelevant Cases.......................... 16
   B. Content Analysis of Relevant Cases (n = 64) ... 19
      1. Quantitative Analyses ............................... 20
         a. Variables ............................................. 20
            Table 4: Study Variables......................... 20
         b. Analytic Strategy .................................. 22
      2. Qualitative Content Analysis .................... 23
         a. Phase One: Preliminary Protocol
            Development.......................................... 23
         b. Phase Two: Case Classification and Inter-
            rater Reliability.................................... 23
         c. Phase Three: Case Studies...................... 24

III. RESULTS ................................................................. 24
   A. Quantitative Findings ................................. 24
      1. Nature of Case Inquiry .............................. 24
         Table 5: Categories of Cases..................... 24
         Table 6: Crosstabulation of Case Inquiry Type
         and Final Case Outcome ......................... 25
      2. Factual Predicates to CLUV ....................... 26
         Table 7: CLUV Circumstances in the Sample 26
      3. CLUV-Specific Laws .................. 26
      4. Duration of CLUV Incident ..................... 27
      5. Child Age ............................................. 27
         Table 8: Crosstabulation of Age and Final Case
         Outcome.................................................. 27
      6. Prediction of Case Outcomes ................... 27
         Table 9: Logistic Regression of CLUV Case
         Outcomes.................................................. 28
INTRODUCTION

On the morning of June 18, 2014, Justin Ross Harris buckled his twenty-two-month-old son, Cooper, into a rear-facing car seat and asked, “[a]re you ready to go to school?”¹ After spending all day at work, Harris made the discovery that no parent ever wants to make.² Harris never dropped off Cooper at his daycare facility; and as a result, the child died after spending the entire day trapped in the vehicle during the height of Georgia’s scorching summer heat.³ Justin Ross Harris was sentenced to life in prison without the

² Id.
³ See Willingham & Blau, supra note 1.
possibility of parole for the death of Cooper Harris. In June of 2022, the Georgia Supreme Court reversed his conviction on the grounds that unfairly prejudicial information had been admitted into evidence at his trial that should have been excluded.

The death of Cooper Harris quickly gained traction on social media and in the public sphere for a variety of reasons, two of which seem particularly salient. First, the way that law enforcement initially reported on this case fueled speculation that Harris had intentionally left Cooper in the car. Second, parents nationwide condemned Harris, claiming that they would never forget that their child was in the backseat of a vehicle.

Harris routinely frequented the Chick-fil-A near his work after dropping off Cooper at daycare. The day of Cooper’s death, Harris took Cooper out to Chick-fil-A for breakfast as a treat, thus altering his usual morning routine. Changes to a normal routine and the general strains and exhaustion relating to caring for a young child can result in the lack of awareness of the child which has been dubbed by medical researchers as “Forgotten Baby Syndrome.”

---
5 Harris v. State, 875 S.E.2d 659, 665–66 (Ga. 2022). The wrongfully admitted evidence included information about Harris’ sexual activities as circumstantial “bad character” evidence of his motive to have intentionally killed his son. Id. at 685–87, 693–94.
8 Harris, 875 S.E.2d at 669.
9 Id.
Harris may have been continuing his usual morning routine of proceeding straight to work after going to Chick-fil-A, despite having just buckled Cooper into the rear-facing car seat. Harris’ actions suggest that he genuinely thought he had brought Cooper to daycare, as evidenced by the fact that he texted his wife that afternoon to ask what time she planned to retrieve Cooper from daycare. Prosecutors, however, argued that this justification was part of a ruse to make the death seem accidental.

Harris’ case demonstrates the complexity of incidents in which children were left unattended in vehicles (“CLUV”). Regardless of whether the child’s caregiver truly forgets a child in a vehicle or makes a conscious decision to leave the child unattended, a fatal result sparks moral outrage. Part I of this Article presents a synthesis of the extant literature on the CLUV phenomenon. The first section in Part I summarizes the dangers of leaving children unattended in vehicles and the second section reports the prevalence and conditions under which this phenomenon occurs. The third section of Part I analyzes the laws that some state legislatures enacted to curb the CLUV phenomenon, focusing on criminal legal responses. And the fourth section explores the ways in which police and prosecutors typically act in such cases. The balance of the Article presents an original, empirical content analysis of how courts adjudicate CLUV cases. Importantly, our study is not limited to pediatric hyperthermia cases in which a child died. Rather, we examine the complete spectrum of CLUV cases, including those in which there were no fatalities. Part II presents the research methodology we used to conduct the study. Part III presents our results. And Part IV concludes with a discussion on the overall effectiveness of these laws and alternative ways to address children being harmed while unattended in vehicles.

I. REVIEW OF THE LITERATURE

Changes in vehicle technology, especially during the mid-1990s, increased the prevalence of CLUV. For instance, in an effort to make cars safer, airbags became standard equipment in

12 See id.
14 See id.
Children Left Unattended in Vehicles

vehicles during the 1990s. This effort undoubtedly saved the lives of many adult passengers, but it was a common source of fatalities for young children, particularly infants. Throughout nationwide campaigns in the mid-1990s, trends of front seat passenger fatalities began to decline for infants and younger children as more parents took the recommendation to have their children sit in the back seats, preferably in a car seat. Paradoxically, as airbag fatalities for children began to decline, the number of children dying from vehicular heatstroke increased during this time period. As law professor Erika Breitfeld explained, “[t]his [seating change] created a new danger because parents could no longer see their children while driving or exiting the vehicle. Amplifying the problem, children frequently fall asleep during car rides and thus remove potential triggers that indicate their presence, such as crying, cooing, babbling, or talking.”

A. Dangers of CLUV

The National Highway Transportation Safety Administration estimates that leaving children in a vehicle is one of the leading causes of vehicle death not associated with a crash. In fact, an average of thirty-eight children under the age of fifteen die of heatstroke in CLUV incidents each year. Given the potentially tragic outcome attendant to CLUV, multiple disciplines—including

17 See Nichols et al., supra note 16, at 317.
medicine and climatology—have come together to better understand what exactly happens when this phenomenon occurs.22

The first of the major concerns attendant to CLUV is the weather conditions surrounding the incident.23 Johannes Horak, an Austrian climatologist, and his colleagues created a model using the outdoor ambient temperature, the thickness of the car windshield and glass, as well as the wind levels outside the car.24 These researchers also manipulated characteristics of the car such as the color, amount of insulation, and the amount the windows were left open allowing for airflow.25 All conditions resulted in an expected increase in temperature consistent with what is known about trapping heat from solar radiation; but the condition with the windows rolled down took roughly double the time to reach maximum temperature.26 Catherine McLaren, a medical scientist from Stanford, and her colleagues evaluated the effects of different levels of outdoor temperatures ranging from seventy-two degrees to ninety-six degrees Fahrenheit.27 These researchers found that regardless of the initial temperature, the internal temperature rate change was the same.28 These studies challenge two of the most common misconceptions about CLUV: (1) that opening a window will make the interior temperature more tolerable; and (2) that engaging in the conduct is only dangerous when ambient outside air temperatures are close to extremes.29

Infants are particularly vulnerable to heat-related deaths because the amount of surface area required to properly regulate

24 See Horak et al., supra note 23, at 108–09.
25 See id. at 109–11.
26 See id. at 112.
27 See McLaren et al., supra note 23, at 109.
28 Id. at 110.
temperature is not fully developed.\textsuperscript{30} To demonstrate this issue, researchers created a model using information from infant vehicular heat stroke cases in Texas in which parents had forgotten a baby in a vehicle while on their way to work.\textsuperscript{31} Despite variation in morning temperatures and solar radiation across the four seasons, when a one-year-old infant was left in a vehicle at 8:00 AM, death occurred no later than 2:00 PM in winter and as early as 10:05 AM in summer.\textsuperscript{32} Even when an infant was not left unattended at the start of a workday, if a vehicle was exposed to direct sunlight, infants could begin to suffer health damage in as little as five minutes and die of heat stroke within an hour.\textsuperscript{33} These findings illustrate the clear need to increase awareness of the dangers of CLUV and to also find ways of preventing it from happening.

\section*{B. Prevalence and Characteristics of CLUV Incidents}

Much of what is known about pediatric vehicular heatstroke cases comes from the organization “No Heat Stroke.”\textsuperscript{34} Researcher and meteorologist Jan Null is the founder of the organization and has compiled a robust database with extensive details documenting over 920 cases of pediatric vehicular heat stroke

\begin{itemize}
\item \textsuperscript{30} See Pietro Ferrara et al., \textit{Children Left Unattended in Parked Vehicles: A Focus on Recent Italian Cases and a Review of Literature}, \textit{39 Italian J. Pediatrics} 71, 71 (2013). A local news station in Tampa, Florida interviewed physician Tiffany Hernandez of the Pediatric Health Care Alliance who explained that children under the age of four are at high risk because their bodies heat up at a rate three to five times faster than adults. Specifically, “[a] child’s thermoregulatory system is not fully developed so they absorb more heat and are less able to lower their body temperature by sweating. When a child’s internal temperature is 104 degrees, their organs start shutting down. At 107 degrees, they could die.” \textit{Warning: Kids Heat Up Faster than Adults in Cars}, WTSP NEWS (Aug. 5, 2016, 10:51 PM), http://www.wtsp.com/article/news/health/warning-kids-heat-up-faster-than-adults-in-cars/67-289255696 [http://perma.cc/CQL7-2CHU].
\item \textsuperscript{32} Id.
\item \textsuperscript{33} See Andrew J. Grundstein, Sara V. Duzinski & Jan Null, \textit{Impact of Dangerous Microclimate Conditions Within an Enclosed Vehicle on Pediatric Thermoregulation}, \textit{127 Theoretical Applied Climatology} 103, 103 (2015) [hereinafter \textit{Impact of Dangerous Microclimate Conditions}].
\end{itemize}

Using 11 different starting cabin air temperatures, we modeled the length of time for a child to reach two critical thresholds: uncompensable heating and heatstroke under “worst case” scenarios. All simulations used a starting dew point temperature of 20°C, and the assumption that all perspiration was evaporated into the air . . . . Under all scenarios, uncompensable heating occurred within 10 min and in most cases within 5 min indicating that the child is no longer capable of balancing the incoming sources of energy and his core body temperature begins to rise monotonically. Thus, very shortly after entering the car, the child is exposed to a microclimatic environment which makes maintaining homeostasis difficult.

\textit{Id.} at 105.

\begin{itemize}
\item \textsuperscript{34} See Null, \textit{No Heat Stroke}, supra note 22.
\end{itemize}
since 1998. Null obtained information relating to the factors surrounding each child’s death, including the reason the child was left unattended, the temperature, and the length of time. This database helps provide an understanding of what these incidents look like to better evaluate the current criminal-legal response. Table 1 presents the frequency of the first 907 of these deaths (covering the twenty-three-year period between 1998 and 2021) according to the circumstances under which they occurred.

<table>
<thead>
<tr>
<th>Circumstance</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Forgotten in Vehicle</td>
<td>477 (52.59%)</td>
</tr>
<tr>
<td>Child Gained Access to Vehicle</td>
<td>234 (25.79%)</td>
</tr>
<tr>
<td>Child Knowingly Left in Vehicle</td>
<td>182 (20.07%)</td>
</tr>
<tr>
<td>Unknown Circumstances</td>
<td>13 (1.54%)</td>
</tr>
<tr>
<td>Total</td>
<td>907 (100%)</td>
</tr>
</tbody>
</table>

Approximately 38% of the total number of CLUV fatalities reported in Table 1 stem from the actions of mothers, whereas fathers account for 25% of such deaths; the remainder are attributable to the actions of both parents, other relatives, or childcare providers. This distribution may be due to the fact that mothers knowingly leave children unattended in vehicles at a rate more than three times that of fathers (59% compared to 18%, respectively). Fathers, however, are responsible for the highest percentage of pediatric heatstroke fatalities stemming from CLUV incidents in which children are “forgotten” (33%), while mothers account for 28% of such deaths. Of the cases in

35 See id.
37 Null, No Heat Stroke, supra note 22; NULL, HEATSTROKE DEATHS: BY THE NUMBERS, supra note 36 at 9 fig.6.
38 NULL, HEATSTROKE DEATHS: BY THE NUMBERS, supra note 36 at 21 fig.9.
39 Id. at 24 fig.9c. It is important to note that of the roughly one-fifth of cases in which a caregiver knowingly left a child unattended in a vehicle, the vast majority do not involve any malicious intent to cause harm to the child, but rather involve caregivers being unaware of the risks summarized in Part I.A. See Washabaugh, supra note 7, at 199; see also KIDS & CARS SAFETY, supra note 18, at 15 (“The overwhelming majority of hot car deaths do NOT involve abuse, neglect, prior history with CPS, drugs or alcohol.”) (emphasis in original).
40 See NULL, HEATSTROKE DEATHS: BY THE NUMBERS, supra note 36 at 22 fig.9a. (Although beyond the scope of this Article, we note that these percentages align with traditional gender roles in which mothers are the predominant caretakers.) Both mothers and fathers succumb to “Forgotten Baby Syndrome.” See generally David Diamond,
which a child was forgotten in a vehicle by their caregiver and
died of heatstroke, 46% were the result of a caregiver not
dropping off the child at a daycare facility.41

The average age of pediatric vehicular heatstroke victims is
27.2 months.42 As Table 2 illustrates, more than half of all such
deaths (n = 490, 54.02%) involve a child who was one-year-old or
younger.43

<table>
<thead>
<tr>
<th>Age</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less Than One Year</td>
<td>278 (30.65%)</td>
</tr>
<tr>
<td>One Year Old</td>
<td>212 (23.37%)</td>
</tr>
<tr>
<td>Two Years Old</td>
<td>171 (18.85%)</td>
</tr>
<tr>
<td>Three Years Old</td>
<td>130 (14.33%)</td>
</tr>
<tr>
<td>Four Years Old</td>
<td>55 (6.06%)</td>
</tr>
<tr>
<td>Five Years Old</td>
<td>28 (3.09%)</td>
</tr>
<tr>
<td>Six Years Old</td>
<td>9 (0.99%)</td>
</tr>
<tr>
<td>Seven through Fourteen Years Old</td>
<td>22 (2.65%)</td>
</tr>
</tbody>
</table>

Importantly, the frequency of fatal pediatric vehicular
heatstroke decreases as the children’s age increases, lending
credence to the finding that most of these deaths are accidental as
a result of children who may be sleeping being forgotten in the
back seat. In fact, of the cases in which a caregiver forgot a child
was in a vehicle, 409 (85.74%) involved a child aged two or
younger.44 Children in that same age range also account for the
largest proportion of pediatric vehicular heatstroke fatalities (n =

Cognitive and Neurobiological Perspectives on Why Parents Lose Awareness of Children in Cars 1 (Aug. 9, 2018), http://www.usf.edu/arts-sciences/departments/psychology/documents/david-diamond-research-on-why-parents-forget-children-in-hot-cars.pdf [http://perma.cc/3MVW-B4HB] (hypothesizing how parents forget children in cars: “[1] the driver loses awareness of the presence of the child in the car; 2) the driver exhibits a failure of the brain’s ‘prospective memory’ system; 3) intervening events during the drive, including stressors and strong distractions, may contribute to the cause of the failure of ‘prospective memory’”); Breitfeld, supra note 19, at 78–84 (summarizing memory failures that can occur when people juggle child care and work responsibilities). Forgetting is more prevalent among fathers, whereas mothers knowingly leave children unattended in vehicles more frequently. See NULL, HEATSTROKE DEATHS: BY THE NUMBERS, supra note 36, at 22 fig.9a, 24 fig.9e. These disparities might be a function of the fact that mothers run more errands while taking care of children, and transportation of children on workdays may be outside the scope of fathers’ regular routines who forget children while their brains are on “autopilot.” See Breitfeld, supra note 19, at 78, 83 (quoting Skip Hollandsworth, The Utterly Heartbreaking and Horrifying Hot-Car Death of Baby Fern Thedford, Tex. MONTHLY (Aug. 23, 2018), http://www.texasmonthly.com/news/hot-car-death-children-michael-thedford-texas/ [http://perma.cc/HUW5-CA92]).

41 See NULL, HEATSTROKE DEATHS: BY THE NUMBERS, supra note 36, at 9 fig.6.
42 Id. at 17 fig.8.
43 NULL, NO HEAT STROKE, supra note 22.
44 See NULL, HEATSTROKE DEATHS: BY THE NUMBERS, supra note 36, at 18 fig.8a.
139, 76.37%) that resulted from a caregiver knowingly leaving a child unattended in a vehicle.\footnote{See id. at 20 fig.8c.} The majority of discoveries of pediatric vehicular heatstroke deaths occur at home (n = 515, 56.78%), followed by discoveries at work (n = 210, 23.15%) and at child care locations (n = 65, 7.17%).\footnote{See id. at 25 fig.10.}

These descriptive statistics hint at the emotional and intellectual strains associated with parenting during a child’s first few years of age.\footnote{See Washabaugh, supra note 7, at 200–01 (first citing The Myth of Joyful Parenthood, ASS’N FOR PSYCH. SCI (Jan. 31, 2011), http://www.psychologicalscience.org/news/were-only-human/the-myth-of-joyful-parenthood.html [http://perma.cc/9V4P-HPCV]; then citing Leslie Irish Evans, Parenthood Is Hard and Scary, HUFFPOST (Nov. 28, 2012), http://www.huffpost.com/entry/parenthood_b_1923288 [http://perma.cc/96XX-JYEX]; and then citing Alice G. Walton, How to Enjoy the Often Exhausting, Depressing Role of Parenthood, THE ATLANTIC: HEALTH (Jan. 9, 2012), http://www.theatlantic.com/health/archive/2012/01/how-to-enjoy-the-often-exhausting-depressing-role-ofparenthood/250901 [http://perma.cc/G2A2-J6KG]).} Especially in the early stages of raising a child, parents may be more sleep deprived and emotionally drained, resulting in an increased likelihood of them forgetting their child in the vehicle.\footnote{See id. at 200–01.} Alternatively, if a parent is out running errands and the newborn child is peacefully sleeping in the car seat, letting the child rest instead of going through the process of waking the child up may be seen as the better option for sleep deprived and emotionally drained parents.\footnote{See id. at 199 (citing Kim Brooks, I Left My Son Alone in the Car for Five Minutes—and It Caused a Two-Year Legal Nightmare, GOOD HOUSEKEEPING (Aug. 14, 2018), http://www.goodhousekeeping.com/life/parenting/a22724843/kim-brooks-son-legal-battle [http://perma.cc/25NE-RRQK]).} Lastly, the ages at which a child gains access to the car on their own is largely self-explanatory. Generally, as children grasp the concept of walking, they become increasingly more difficult to keep track of, thus leading to the potential risk of the child ending up in a vehicle on their own.

C. CLUV-Specific Laws

Currently, twenty-one states have enacted statutes specifically relating to leaving a child in a vehicle.\footnote{See Breitfeld, supra note 19, at 101 (“[A]t least twenty-one states have laws against leaving children unattended in a vehicle.”); see also Jan Null, Unattended Child in Vehicle Laws, NO HEAT STROKE (Feb. 2018), http://www.noheatstroke.org/Laws.pdf [http://perma.cc/X8BC-AJX8].} Some of these state laws were enacted in response to a pediatric hyperthermia vehicle fatality. For example, in 2001, California enacted Senate Bill 255, more commonly known as Kaitlyn’s Law.\footnote{2001 Cal. Stat. ch. 855 § 2 (codified as amended at CAL. VEH. CODE § 15620 (West 2003)).} A babysitter left six-
month-old Kaitlyn Russell in the backseat of a car.52 When she was discovered, the interior temperature of the vehicle exceeded 130 degrees Fahrenheit.53 The babysitter was convicted of involuntary manslaughter and spent ninety days in county jail.54

Kaitlyn’s Law’s provides:

(a) A parent, legal guardian, or other person responsible for a child who is [six] years of age or younger may not leave that child inside a motor vehicle without being subject to the supervision of a person who is [twelve] years of age or older, under either of the following circumstances:

(1) Where there are conditions that present a significant risk to the child’s health or safety.

(2) When the vehicle’s engine is running or the vehicle’s keys are in the ignition, or both.55

Kaitlyn’s Law made it illegal for any individual directly responsible for the care of a child under the age of six to leave that child unsupervised in a vehicle.56

Violations of Kaitlyn’s Law are punishable by a $100 fine and potentially mandated participation in an education program about the dangers of CLUV.57 The law specifically provides that nothing in it “shall preclude prosecution under . . . any other provision of law.”58 As a provision in the vehicle code, Kaitlyn’s Law does not require proof of mens rea for conviction; it is a strict liability public health, safety, and welfare offense.59 Several other states enacted similar strict liability offenses to curtail CLUV,60 These laws arguably help to raise awareness about the dangers of unattended children in vehicles, as well as parents leaving young children in the care of other minors not mature enough to understand the

53 Id.
55 CAL. VEH. CODE § 15620(a)(1)–(2) (West 2003).
56 Id.
57 See id. § 15620(b).
58 Id. § 15620(c).
risks that heat poses to particularly young children. Thus, these laws may help reduce the prevalence of people intentionally leaving children unattended. But these laws cannot deter the phenomenon of accidentally forgetting about a child in a vehicle “which can happen to anyone at any time.”\textsuperscript{61}

In contrast to strict liability approaches to CLUV, some states require proof of mens rea, usually at the level of recklessness,\textsuperscript{62} although some laws impose liability for negligence, while others reserve sanctions for intentional acts.\textsuperscript{63} Nevada is the only state to exempt unintentional actions from liability.\textsuperscript{64}

These laws also vary with regard to liability being predicated on a minimum period of time, such as five or fifteen minutes.\textsuperscript{65} Other CLUV laws are more ambiguous, only sanctioning the conduct if it occurs for a period of time that poses a substantial risk to the wellbeing of the child.\textsuperscript{66} In our opinion, such approaches are flawed. These laws signal that it is acceptable to leave children unattended in vehicles so long as it is just for a specific period of time or if the child is younger than a particular age. But these laws should communicate that CLUV is not acceptable for children of any age or for any duration of time, rather than telegraphing it may be acceptable under the right circumstances.

Are such laws the best way to address CLUV? Consider the case of Brittany Borgess. Borgess was sleep deprived and encountered unexpected road construction that altered her normal route to work.\textsuperscript{67} Instead of taking her regular exit for her daughter’s daycare, Borgess proceeded straight to work and did not realize until after the end of the day that her daughter was still in the car.\textsuperscript{68} Despite being incredibly distraught over the error that claimed the life of her four-year-old, she was charged

\textsuperscript{61} Id. at 207.
\textsuperscript{62} See id. (citing CONN. GEN. STAT. ANN. § 53-21a (West 2012); MICH. COMP. LAWS ANN. § 750.135a (West 2009); OKLA. STAT. ANN. tit. 43A, § 10-103 (West 2019); 75 PA. STAT. AND CONS. STAT. ANN. § 3701.1 (West 2006)).
\textsuperscript{63} See id. at 208 (first citing UTAH CODE ANN. § 76-10-2202 (West 2011) (permitting punishment upon a showing of criminal negligence or any higher level of mens rea); and then citing 720 ILL. COMP. STAT. ANN. 5/12C-5 (West 2013) (requiring purpose or knowledge for liability)).
\textsuperscript{64} See id. (citing NEV. REV. STAT. ANN. § 202.485 (West 2017)).
\textsuperscript{65} See id. at 205 (first citing FLA. STAT. ANN. § 316.6135 (West 2014) (fifteen minutes); and then citing TEX. PENAL CODE § 22.10(a) (West 2021) (five minutes)).
\textsuperscript{66} See id. (citing CONN. GEN. STAT. ANN. § 53-21a (West 2012); MICH. COMP. LAWS ANN. § 750.135a (West 2009)).
\textsuperscript{68} Id.
with involuntary manslaughter and child endangerment. Because the law in Florida required proof of recklessness—conscience disregard of a known risk—she was acquitted of these offenses, but convicted of the lesser strict liability charge of leaving an unattended child in a vehicle, a summary offense for which a $25 fine was imposed. When considering the mental health toll associated with processing the loss of a child—especially when due to one's own mistake—subsequent criminal prosecution undoubtedly exacerbates an already incredibly difficult situation.

D. Charging and Prosecuting Decisions in Pediatric Hyperthermia Death Cases

Of course, both police and prosecutorial discretion impact how all CLUV cases are handled. Prior research has focused on the exercise of that discretion in cases resulting in a child's death. For instance, the organization “Kids and Car Safety” constructed a database of lethal pediatric hyperthermia cases similar to No Heat Stroke's, but the former’s data includes information about how the criminal legal system responded to such cases. Table 3 summarizes the case outcomes they tracked across the thirty-year period between 1990 and 2020.

The organization noted that of the 31% of cases in which a conviction was ultimately obtained, many resulted from defendants entering a plea to avoid the re-traumatization that would likely occur at trial while they were trying to cope with “the tragic loss of child.” It also broke down outcomes based on the context of incidents. In the “forgotten” child cases (i.e., when CLUV occurs unknowingly), 41% did not result in charges; whereas of the cases charged, 32% resulted in a conviction and 11% resulted in an acquittal. When children gained access to vehicles on their own volition, the rate of charge declinations skyrocketed to 75%, while 9% resulted in convictions and 45 ended in acquittal. By contrast, in cases involving conscious decisions to leave a child unattended in vehicles, prosecutors filed charges in 84% of cases, 69% of which ended in convictions compared to just 6% ending in acquittals.

---

69 See id.
71 See Null, NO HEAT STROKE, supra note 22.
72 See KIDS & CAR SAFETY, supra note 18, at 15.
73 See id. at 15 fig.12.
74 See id. at 15.
75 Id. at 16 fig. 12a.
76 Id. at 16 fig.12b.
77 Id. at 17 fig.12c.
II. METHODS

As previously mentioned, the outcomes presented in Table 3 are limited to cases in which a child died. Incidents in which a child survives do not receive the same level of media or scholarly attention. As a result, we know little about the criminal legal system response to nonlethal CLUV cases. To fill this gap in the literature, the researchers searched Westlaw’s state caselaw database using the following Boolean parameters for cases decided in the thirty-one-year period between January 1, 1990, and December 31, 2021:

unattend! /s child! /s (car or vehicle or truck)

The results were filtered to exclude civil cases but included both published and unpublished criminal cases. This search resulted in a sampling frame of 185 cases.

A. Removal of Irrelevant Cases

Of the 185 total cases, we excluded 121 of them because they were not relevant to the research question. These cases were excluded for one or more of the following four reasons.

First, we removed eight duplicate cases. Duplicate cases typically occurred when a lower court rendered a decision that was then appealed or when defendants filed successive petitions for post-conviction relief. In such cases, the final decision on the merits is included in the search sample.78

Second, the search terms pulled fifty-two cases that referenced the words “child” and “vehicle,” but under

---

circumstances having nothing to do with CLUV. 79 Third, despite selecting the criminal case filter in Westlaw, the search results

included forty-four cases from family court and five other types of judicial proceedings.

Finally, we removed twelve cases from our final sample cases in which a child had been left unattended in a vehicle but that fact had nothing to do with the central issue in the judicial decision.


B. Content Analysis of Relevant Cases (n = 64)

“Content analysis is a formal system for doing something we all do informally rather frequently—draw conclusions from observations of content.”83 But unlike with informal observations, content analysis employs systematic procedures “for making replicable and valid inferences from [data] . . . to the contexts.”84

We used a mixed-methods approach for the present study by conducting both quantitative and ethnographic content analyses.85 The quantitative portion of our content analysis applies an a priori design to review both published and unpublished judicial opinions in criminal cases and code for the presence or absence of predefined

*1 (Tenn. Crim. App. Oct. 8, 2001), involved an appeal of a murder conviction. One of the key pieces of evidence tying the defendant to the homicide involved an eyewitness who identified the defendant as a function of a CLUV.

Just prior to the stabbing incident, . . . the Barretts were returning from a shopping trip when Mr. Barrett noticed an African-American infant asleep in a vehicle as they walked through the apartment parking lot. The baby was strapped into a child’s car seat located on the back seat of the vehicle, the windows were down, and the vehicle was unattended. The Barretts were concerned for the baby’s safety. Since they could see the parking lot from their apartment, they went home and kept a watch on the car from their window. After twenty or thirty minutes passed and no one came to check on the baby, Barrett wrote down the license plate number of the car, and Mrs. Barrett called 911 to report the possibility that someone may have abandoned a child. Approximately one hour after Barrett first noticed the infant, he observed two men walk hurriedly across the parking lot and get into the car with the baby in it. It was dark and he could not see their faces, but he was able to determine that they were African American. Barrett described the taller man as extremely “hyper” and “on the move.” The shorter, husky man seemed calmer and “in control.” The husky man drove the vehicle; he backed up slowly, and then headed for the exit with the headlights off. When the men reached the road, the headlights came on and the vehicle began to accelerate.


83 See Guido H. Stempel, III, Content Analysis, in MASS COMMUNICATION RESEARCH AND THEORY 209 (Guido H. Stempel et al., 2003).
84 See Klaus Krippendorff, Content Analysis: An Introduction to Its Methodology 18 (Margaret H. Seawell et al. eds., 2d ed. 1980).
85 See David L. Altheide & Christopher J. Schneider, Qualitative Media Analysis 24–26 (Vicki Knight et al. eds., 2d ed. 2013) (detailing the phases of qualitative content analysis); Kimberly A. Neuendorf, The Content Analysis Guidebook 351 (Karen Omer et al., 2d ed. 2017).
variables. This approach allows for the calculation of both the frequency of key variables, as well as the extent to which they may be related. It also allows for hypothesis testing and promotes replicability. The qualitative part of content analysis is inductive. This method is particularly well-suited for comparing and contrasting multiple cases (like the sixty-four relevant cases in our research sample) to detect emergent themes across the cases.

1. Quantitative Analyses

After both researchers read and agreed on the removal of irrelevant cases, sixty-four relevant cases remained. All sixty-four cases in the final research sample centered around criminal or quasi-criminal liability for leaving a child unattended in a vehicle.

a. Variables

Each judicial opinion was coded for manifest content for all the variables presented in Table 4.

<table>
<thead>
<tr>
<th>VARIABLE NAME</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td></td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td>A dichotomous nominal-level variable indicating final resolution of a claim in favor of (1) prosecution (defendant convicted/conviction upheld), or (2) defense (defendant acquitted/conviction reversed either on appeal or collateral attack).</td>
</tr>
<tr>
<td>Independent Variables</td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>A multicategory ordinal-level variable classifying the child victim’s age as (1) infant (zero to twenty-three months); (2) toddler (two to three years); (3) early childhood (four to six years); (4) school-age (seven to twelve years); or (5) teenager (thirteen to eighteen years).</td>
</tr>
</tbody>
</table>

---

86 See Neuendorf, supra note 85, at 18.
87 See Altheide & Schneider, supra note 85, at 96–119 (discussing variable coding); id. at 24–26 (discussing variable measurement).
88 See id. at 24–26.
89 See id. at 27.
90 A nominal variable is sometimes referred to as a categorical variable because it expresses categories that have “no intrinsic value.” What Is the Difference Between Categorical, Ordinal and Interval Variables? UCLA Advanced Rsch. Computing, http://stats.oarc.ucla.edu/other/mult_pkg/whatstat/what-is-the-difference-between-categorical-ordinal-and-interval-variables/ [http://perma.cc/P3WV-VMSV] (last visited Feb. 25, 2023). They can be dichotomous, such as yes/no; guilty/not guilty, or they can have numerous categories, such as a range of hair colors or a listing of races or ethnicities. Id.
91 An ordinal-level variable is similar to a nominal one, but there is a clear ordering of the categories, such as expressing height as being short, average, or tall; or expressing income as low, medium, and high. Id.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circumstances</td>
<td>A trichotomous, nominal-level variable indicating whether (1) a caretaker intentionally left a child in a vehicle, (2) a caretaker forgot about a child in vehicle, or (3) a child accessed a vehicle without a caretaker's knowledge.</td>
</tr>
<tr>
<td>Charge</td>
<td>A multicategory, nominal-level variable for which dummy codes were created to indicate whether the defendant had been charged with (1) child endangerment or similar charge requiring recklessness; (2) child abuse; (3) child neglect; (4) an intentional homicide requiring knowledge or purpose; (5) an unintentional homicide requiring recklessness; or (6) a criminally negligent homicide.</td>
</tr>
<tr>
<td>CLUV Law</td>
<td>A dichotomous, nominal-level variable indicating whether the state in which the case arose had a specific statutory provision governing leaving children unattended in vehicles.</td>
</tr>
<tr>
<td>Temperature</td>
<td>A dichotomous, nominal-level variable that was dummy coded to indicate that the temperature was either (1) hot (e.g., summer), or (2) cold (e.g., winter).</td>
</tr>
<tr>
<td>Region</td>
<td>A multicategory, nominal-level variable for which dummy codes were created to indicate the geographical region of the country based on U.S. Census regions: (1) Northeast, (2) Midwest, (3) South, and (4) West.</td>
</tr>
<tr>
<td>Time</td>
<td>A multicategory, ordinal-level variable indicating how long a child was left unattended in a vehicle coded as (1) short (five to thirty minutes), (2) medium (thirty-one to sixty minutes), (3) long (sixty-one minutes or longer).</td>
</tr>
</tbody>
</table>

92 Dummy variables express how nominal variables are coded for regression analyses. See notes 98–100 and accompanying text. For example, when coding the nominal variable of hair color used as an example in the preceding footnote, “blond” might be dummy coded with a value of 1, “brunette” with a value of 2, “red” with a value of 3, and “other” with a value of 4. Because the values assigned to these categories are devoid of intrinsic value, the common parlance in statistics is to refer to them as being “dummy” coded. See, e.g., How to Use Dummy variables in Regression Analysis, STATOLOGY (Feb. 1, 2021), http://www.statology.org/dummy-variables-regression/ [http://perma.cc/WYA6-PULG]; see generally MELISA A. HARDY, REGRESSION WITH DUMMY VARIABLES 7–17 (1993) (“creating dummy variables” chapter); id. at 18–28 (“using dummy variables as regressors” chapter). In the present study, because the charges in any given case are nominal variables, they are dummy coded for analysis.


94 This region includes Iowa, Ill., Ind., Kan., Mich., Minn., Mo., Neb., N.D., Ohio, S.D., and Wis. Id.

95 This region includes Ala., Ark., Del., D.C., Fla., Ga., Ky., La., Md., Miss., N.C., Okla., S.C., Tenn., Tex., Va., and W. Va. Id.

96 This region includes Alaska, Ariz., Cal., Colo., Haw., Idaho, Mont., Nev., N.M., Or., OR, Wash., and Wyo. Id.

97 Time was originally measured as a ratio-level variables (i.e., actual time in minutes). Given the wide range of values, however, we collapsed time into the three categories specified in Table 4 to facilitate regression analysis without extreme outliers that, for the purposes of hyperthermia, are not relevant for the reasons explained in Part I.A. See supra notes 23–33 and accompanying text.
b. Analytic Strategy

The analytic strategy we used to examine how the criminal-legal system approaches CLUV cases involved a two-step sequence. First, we ran descriptive statistics to provide an overview of the data. Second, we ran a series of inferential statistics—including chi-squares, Fisher Exact tests, and logistic regressions—to examine the effects and predictive probability of the independent variables listed in Table 4 on case outcomes.

Chi-square tests and Fisher Exact statistics determine whether there are significant differences between nominal (i.e., categorical) variables. Put simply, these tests are used to see if there is a relationship between two seemingly unrelated variables. Regression is a statistical procedure that assesses whether a set of independent variables is associated with some outcome, referred to as the dependent variable. These analyses not only identify the particular independent variables that significantly predict the outcome, but also the degree to which they do so as “indicated by the magnitude and sign of the beta estimates.” There are several types of regression, including logistic regression, which is the appropriate type when the dependent variable is dichotomous, meaning there are only two possible outcomes. The present study is dichotomous, with the final case outcome in favor of either the prosecution or the defense. The final regression model reported in Table 9 only uses the amount of time a child was left unattended in a vehicle and three independent variables that were statistically

98 See, e.g., ALAN AGRESTI, STATISTICAL METHODS FOR THE SOCIAL SCIENCES § 8.2 (Suzy Bainbridge et al. eds., 5th ed. 2018); DAVID WEISBURD & CHESTER BRITT, STATISTICS IN CRIMINAL JUSTICE 197–233 (4th ed. 2014). The test is instrumental in determining the independence of or relationship between cross-tabulated data. The statistical procedure tests whether an association exists between the two variables by comparing the observed pattern of responses in the cells to the pattern that would be expected if the variables were truly independent of each other. Calculating the chi-square statistic and comparing it against a critical value from the chi-square distribution allows the researcher to assess whether the observed cell counts are significantly different from the expected cell counts. See Using Chi-Square Statistic in Research, STAT. SOLS. http://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/using-chi-square-statistic-in-research/ [http://perma.cc/2A6X-F3AN] (last visited Aug. 3, 2021).

The Fisher Exact test serves the same purpose as chi-squares but is more appropriate with sample sizes less than 1,000 (like ours) or when 20% of expected frequencies in cross-tabulation cells are less than or equal to 5. See Matthias Döring, Testing Independence: Chi-Squared vs Fisher’s Exact Test, DATA SCI.BLOG (Oct. 17, 2018), http://www.datascienceblog.net/post/statistical_test/contingency_table_tests/ [http://perma.cc/C36Q-4WG8]. Because the research sample consists of 64 cases, we report both the chi-square and Fisher’s Exact statistics as confirmatory of each other.


100 Id.

101 Id.; see also JASON W. OSBORNE, BEST PRACTICES IN LOGISTIC REGRESSION 3–4 (2015).
significant \((p < 0.05)\) or approached significance \((p < 0.08)\) in the chi-square analyses; it omits those variables that were insignificant \((p > 0.08)\). As a result, in addition to the duration of time in a CLUV-incident, the logistic regression model includes three other predictor variables: the child’s age \((p < 0.014)\); whether the jurisdiction has a CLUV-specific law \((p < 0.052)\); and type of case inquiry using the categories that emerged in the qualitative analyses of case \((p < .022)\), as summarized in Table 5.

2. Qualitative Content Analysis

The qualitative portion of this study focused on narrative data in which both categorical and unique data were obtained from each case studied. The content analysis was conducted in three phases, the completion of which allowed for the creation of a typology based on the patterns that emerged during the analyses of cases.

a. Phase One: Preliminary Protocol Development

During the first phase, each of the two researchers independently reviewed ten relevant cases to identify distinctive patterns in ways that courts adjudicated cases involving CLUV. This allowed us to develop a preliminary protocol for coding cases. We then compared our assessments to harmonize our coding so that all cases presenting similar themes could be more reliably coded as falling within a particular category.

b. Phase Two: Case Classification and Inter-rater Reliability

The second phase required each of the two researchers to code all sixty-four relevant cases independently. Cases falling within one of the themes identified during the first phase were added to that previously-identified category. We created new categories for cases presenting substantially different CLUV issues, thereby allowing for the emergence of central themes that are summarized in Table 5 and explored in detail in Part III.B of this Article.

During our preliminary coding of all 185 cases in the sampling frame, we agreed on all of the 121 cases that were irrelevant, although we initially coded three of these irrelevant cases as meeting different exclusion reasons. Of the sixty-four relevant cases, we are pleased to report that we achieved an impressively

---

102 See generally ALTHEIDE & SCHNEIDER, supra note 85, at 23–73.
103 See id. at 44–45.
high level of inter-rater reliability at 95.3%.\footnote{This is well above the typical threshold of 75% agreement required in the social sciences and even exceeds the 90% agreement rate threshold in medicine. See, e.g., Stephanie Glen, \textit{Inter-rater Reliability IRR: Definition, Calculation, Stat. How To} (July 17, 2016), http://www.statisticshowto.com/inter-rater-reliability/ [http://perma.cc/RG46-XFSH].} That is a function of the fact that both researchers coded all but four of the cases identically. For the four cases on which we initially disagreed, we talked through our differences of opinion and came to an agreement on the best way to code them.

c. Phase Three: Case Studies

In the final phase of the content analysis, we compared and contrasted the cases within each of the categories that emerged in phase two. This allowed us to make some generalizations not only about the factual circumstances under which CLUV results in criminal or quasi-criminal charges being filed, but also about how courts grapple with CLUV issues.

III. RESULTS

A. Quantitative Findings

Although the methodology drew from cases at all levels of judicial proceedings, nearly all of the cases in the sample came from state appellate or supreme courts.

1. Nature of Case Inquiry

Table 5 presents the primary types of claims raised by the facts in the cases, along with their corresponding frequencies.

<table>
<thead>
<tr>
<th>Major Case Themes</th>
<th>Number of Cases</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consequences of Conviction</td>
<td>09</td>
<td>14.06%</td>
</tr>
<tr>
<td>Sufficiency of the Evidence</td>
<td>49</td>
<td>76.56%</td>
</tr>
<tr>
<td>Statutory Construction</td>
<td>06</td>
<td>09.38%</td>
</tr>
<tr>
<td>Total</td>
<td>64</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

As Table 5 illustrates, a supermajority ($n = 49, 76.6\%$) of these cases involved sufficiency of the evidence claims. This is consistent with national data reporting that sufficiency of the evidence claims is the most common legal issue appellate courts address.\footnote{\textsc{Nicole L. Waters Anne Gallegos, James Green & Martha Rozsi}, \textit{U.S. Dep't of Just.}, NCJ 248874, \textit{Criminal Appeals in State Courts} 1 (2015).} Most of these cases involved appellate review of one of two issues: (1) whether the record evidenced sufficient facts proving that the
circumstances presented reckless conduct (i.e., a significant risk of death or serious injury) or (2) whether the record evidence establishes some other level of mens rea necessary for conviction. The remaining cases in the research sample involved (1) challenges to sentences or other consequences of CLUV convictions \( (n = 9, 14.1\%) \), or (2) questions of statutory interpretation and validity \( (n = 6, 9.4\%) \). The qualitative results reported in Part III, Sections B.2 through B.4 explain how courts grappled with each of these types of claims.\(^\text{106}\)

| TABLE 6: CROSSTABULATION OF CASE INQUIRY TYPE AND FINAL CASE OUTCOME |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|
| **Outcome**                     | **General Case Inquiry** |                |                |                |
|                                 | **Sufficiency of the Evidence** | **Consequences** | **Statutory Construction** | **Total** |
| Proseuction                     | Count | 33     | 9       | 2       | 44                               |
|                                 | % within | 67.3% | 100.0% | 33.33% | 68.75%                           |
| Defense                         | Count | 16     | 0       | 4       | 20                               |
|                                 | % within | 32.6% | 0.00%  | 66.66% | 31.25%                           |
| Total                            | Count | 49     | 9       | 6       | 64                               |
|                                 | % within | 76.56%| 14.06% | 9.38%  | 100%                            |

\( \chi^2 = 7.6388, p = .022; \text{Fisher's exact: } p = .017 \)

As Table 6 illustrates, the type of claims being adjudicated was significantly related to case outcomes. Three points stand out from the data. First, although twice as many cases involving appeals based on the sufficiency of the evidence resulted in the affirmance of convictions, sixteen of forty-nine (32.6%) of such cases resulted in appellate decisions in favor of the defense. This is a notable finding because the reversal rate on appeal for insufficiency of the evidence is typically dramatically lower—8.1% nationwide according to the Bureau of Justice Statistics.\(^\text{107}\) Second, courts affirmed sentences or otherwise denied collateral relief from CLUV-related convictions in all nine (100%) of the cases that raised such claims on appeal. This is notably higher than the 83% sentence affirmation rate national for all types of criminal cases, perhaps due to the emotional punch that many of these cases present.\(^\text{108}\) Finally, defendants won twice as many cases as the prosecution when it came to questions of statutory

---

\(^\text{106}\) See infra notes 137–223 and accompanying text.

\(^\text{107}\) WATERS ET AL., supra note 105, at 6 fig.3.

\(^\text{108}\) Id.
interpretation. The qualitative results presented in Part III.B shed light on these findings.

2. Factual Predicates to CLUV

Table 7 shows the distribution of the circumstances under which CLUV occurred in the sixty-four cases in the research sample. The circumstances were not significantly related to case outcomes.

<table>
<thead>
<tr>
<th>Circumstance</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Forgotten in Vehicle</td>
<td>14 (21.88%)</td>
</tr>
<tr>
<td>Child Gained Access to Vehicle</td>
<td>1 (1.56%)</td>
</tr>
<tr>
<td>Child Knowingly Left Unattended in Vehicle</td>
<td>47 (73.44%)</td>
</tr>
<tr>
<td>Unknown Circumstances</td>
<td>2 (3.13%)</td>
</tr>
<tr>
<td>Total</td>
<td>64 (100.00%)</td>
</tr>
</tbody>
</table>

$x^2 = 0.6318, p > 0.73, \text{n.s.; Fisher's exact } p = 1.00$

Despite the lack of statistically significant differences between these reasons and case outcomes, qualitative analysis revealed a notable finding that is not evident from the statistical results. In twelve of the sixty-one (19.6%) cases in which a caregiver either knowingly left a child ($n = 47$) or forgot a child ($n = 14$) in a vehicle, the event coincided with the caregiver either being under the influence of drugs or alcohol, or engaging in criminal activity like shoplifting. This finding is discussed in detail in Part III.B.1.

3. CLUV-Specific Laws

Twenty-seven cases (42.2%) occurred in states that had a specific CLUV law, whereas the remaining thirty-seven cases (57.8%) transpired in states without such laws. Although the chi-square analysis of the relationship between the presence of such a law and case outcome was not statistically significant at the 0.05 alpha level, ($x^2 = 3.784, p < 0.052$), it closely approached significance and therefore is included in the regression model reported in Table 9. Care should be taken in how to interpret this finding. It is likely due, in part, to how long ago some of the cases were decided. Recall that the research sample covered a thirty-one-year span of time between 1990 and 2021. Because CLUV laws are comparably new, that might explain why a majority of the cases occurred in jurisdictions without a CLUV law, rather than indicating that CLUV laws are related to case outcomes. And most importantly, the data does not support any
interpretation that indicates CLUV laws are related to a reduction in the incidence of the phenomenon.

4. Duration of CLUV Incident

Twenty-two cases (34.4%) involved a child being left unattended in a car for thirty minutes or less. Ten cases (15.6%) involved a child being left anywhere from thirty-one minutes to sixty minutes. Twenty-three (35.9%) cases involved a child being left anywhere from 61 minutes to 720 minutes. The amount of time was unknown for nine cases (14.1%). The relationship between the amount of time a child was left unattended and the final case outcome was not statistically significant ($\chi^2 = 0.205, p < 0.903$). Nonetheless, because the duration of a CLUV incident is medically salient, we included this variable in the logistic regression model with an abundance of caution since it might prove significant when controlling for other factors.

5. Child Age

As Table 8 illustrates, the age of the children left unattended in vehicles is significantly related to case outcome.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Infant</th>
<th>Toddler</th>
<th>Early Childhood</th>
<th>School-Age</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prosecution count</td>
<td>26</td>
<td>2</td>
<td>2</td>
<td>9</td>
<td>39</td>
</tr>
<tr>
<td>% within</td>
<td>66.7%</td>
<td>5.0%</td>
<td>5.0%</td>
<td>23.1%</td>
<td>68.4%</td>
</tr>
<tr>
<td>Defense count</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>% within</td>
<td>27.8%</td>
<td>7.1%</td>
<td>29.4%</td>
<td>38.9%</td>
<td>31.6%</td>
</tr>
<tr>
<td>Total count</td>
<td>31</td>
<td>3</td>
<td>7</td>
<td>16</td>
<td>57</td>
</tr>
<tr>
<td>% within</td>
<td>54.3%</td>
<td>5.3%</td>
<td>12.3%</td>
<td>28.1%</td>
<td>100%</td>
</tr>
</tbody>
</table>

$\chi^2 = 9.0353, p = .022$; Fisher’s exact: $p = .014$

6. Prediction of Case Outcomes

Table 9 presents the results of the logistic regression with the case outcome variable and the predictor variables of age, CLUV law, the transformed time variable, and a condensed version of the case classification variable.\textsuperscript{109} Notably, age

\textsuperscript{109} Because all of the appeals challenging sentences or other consequences of convictions were denied (i.e., 100% resolved in favor of prosecution), that category of case dispute type is a perfect predictor of case outcome in the dataset. That category of case therefore needs to be eliminated from the logistic regression model to avoid the so-called “zero-cells” problem. See Xiao Chen, Phil Ender, Michael Mitchell & Christine Wells, UCLA Statistical Consulting Grp., \textit{Logistic Regression Diagnostics}, in \textit{Stata Web Books Logistic}}
continued to be the only significant factor \((p < .008)\). As the age of the victim increased, the odds of the court siding with the prosecution and upholding a conviction decreased.

### TABLE 9: LOGISTIC REGRESSION OF CLUV CASE OUTCOMES

<table>
<thead>
<tr>
<th>Observations</th>
<th>51</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR Chi2(3)</td>
<td>13.82</td>
</tr>
<tr>
<td>Model (p) - value</td>
<td>0.0079</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.2087</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regressors</th>
<th>Odds Ratio</th>
<th>Standard Error</th>
<th>(p) - value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.50</td>
<td>0.13</td>
<td>0.008</td>
</tr>
<tr>
<td>CLUV Law</td>
<td>0.31</td>
<td>0.23</td>
<td>0.117</td>
</tr>
<tr>
<td>Length of Time</td>
<td>1.59</td>
<td>0.66</td>
<td>0.260</td>
</tr>
<tr>
<td>Statutory Construction Claim</td>
<td>0.20</td>
<td>0.22</td>
<td>0.150</td>
</tr>
<tr>
<td>Constant</td>
<td>7.30</td>
<td>7.63</td>
<td>0.057</td>
</tr>
</tbody>
</table>

### B. Qualitative Findings

1. CLUV, Drugs, Alcohol, and Other Crimes \((n = 13, 20.3\%)\)

The first notable theme that emerged during our qualitative review of cases in the research sample was that caregivers in roughly one out of every five cases were either under the influence of alcohol or other drugs at the time of the CLUV incident or they were arrested for other crimes in addition to those related to CLUV.\(^{110}\) The substance use or other offenses appear to have been causally related to the CLUV event in most, if not all, of these cases.

Drug or alcohol use can lead caregivers to become too impaired to perceive the whereabouts of children or, alternatively, to leave children unattended while they obtain or use the substance(s) in question. Consider what occurred in Shouse v. Commonwealth:

Shouse took a Xanax mid-afternoon, and then dropped her two-year-old son off at her mother's while she went shopping with a friend. At about eight in the evening, she retrieved her son and went to their apartment where she took a second Xanax. A friend stopped by at about 10:30 p.m. and stayed until about 12:30 a.m., when Shouse drove the friend to Jeff Burch’s apartment to obtain marijuana. She then drove Burch to a nearby Waffle House and back to his apartment, where they sat in the car and talked for about an hour. Burch gave her some marijuana, but both claim they did not smoke it at that time. At about 3:00 a.m., Shouse drove to a Thornton’s, bought doughnuts and a drink, and then went home. She got several items out of the car, went inside, and fell asleep. She left her son in the car.

Burch and others tried to contact Shouse until about 3:00 p.m. the next day, when her mother went to the apartment to check on her and the child. Shouse, who appeared startled and confused, did not know where her son was. The grandmother ran to the car where the child was still strapped in his car seat. He was pronounced dead at the scene.111

When officers arrived at Shouse’s apartment, they observed “a number of drugs.”112 These facts likely contributed to the state charging Shouse with wanton murder and criminal abuse.113 The state also charged her with wanton endangerment for having driven the vehicle with her child in the car while she was under the influence.114 A jury convicted her of all three charges and a judge sentenced her to thirty-five years imprisonment.115

People v. Rudell similarly illustrates how a defendant’s use of drugs or alcohol seemingly contributes to prosecutors moving forward with criminal charges in CLUV cases, even when a child is rescued.116 Rudell involved a defendant who left her six-month-old child in a car while she attended a party.117 She drank to excess—so much so that an officer testified she appeared to him as being “extremely intoxicated, a 10 on a scale of 1 to 10.”118 A passerby flagged down police to report the unattended infant and firefighters subsequently removed the crying baby from the vehicle.119 After a bench trial, the court convicted her of endangering the life of a child and public intoxication.120 She was

111 Shouse, 481 S.W.3d at 482.
112 Id.
113 See id.
114 Id.
115 Id. at 483.
117 See id. at 543.
118 Id. at 543.
119 See id. at 542–43.
120 See id. at 542, 544. The child endangerment statute contains a specific provision relevant to CLUV:

(a) A person commits endangering the life or health of a child when he or she
sentenced to six months of probation, parenting classes, and alcohol treatment.\textsuperscript{121}

After substance use, shoplifting by childcare providers was the next most common factual circumstance contributing to CLUV.\textsuperscript{122} Importantly, CLUV while caregivers engage in a crime like shoplifting potentially exposes children to a qualitatively different type of risk than substance abuse as \textit{State v. Spivey} illustrates.\textsuperscript{123}

The defendant in \textit{Spivey} left her six-year-old and three-year-old grandchildren in her car while she shoppedlifted headphones from a department store.\textsuperscript{124} After a security guard detained her, she alerted him to the fact that she had left the children unattended in the parking lot.\textsuperscript{125} The guard notified police, who subsequently found the children “wearing winter coats” while sitting in a locked vehicle with “a plastic bag over the front-passerger window and frost on the windows” while the “outside temperature was approximately 15 degrees\textsuperscript{126} After being convicted of two counts of child endangerment, among other charges, she appealed arguing that the evidence was insufficient to prove “that she recklessly created a substantial risk of harm to her grandchildren.”\textsuperscript{127} In upholding her conviction, the \textit{Spivey} court reasoned that:

\begin{quote}
[She] knowingly took the risk of getting caught and being detained for her actions, resulting in the children being left unattended in the car for a potentially unknown amount of time. Ball took the risk that she could have been put in a room by herself with no way to alert anyone to the children in the car. And, even though she was able to tell the
\end{quote}

\begin{itemize}
\item knowing: (1) causes or permits the life or health of a child under the age of 18 to be endangered; or (2) causes or permits a child to be placed in circumstances that endanger the child's life or health. . . .
\item (b) A trier of fact may infer that a child 6 years of age or younger is unattended if that child is left in a motor vehicle for more than 10 minutes.
\end{itemize}

720 ILL. COMP. STAT. 5/12C-5 (2013).

\textsuperscript{121} See \textit{Rudell}, 78 N.E.3d at 544.
\textsuperscript{123} See \textit{Spivey}, 2021 WL 3234383, at *3.
\textsuperscript{124} \textit{Id}.
\textsuperscript{125} \textit{Id}.
\textsuperscript{126} \textit{Id.} at *1, *2.
\textsuperscript{127} \textit{Id}. The child endangerment statute under which the defendant was charged did not contain a specific CLUV provision, but rather provides as follows:

No person, who is the parent, guardian, custodian, person having custody or control, or person in loco parentis of a child under eighteen years of age or a mentally or physically handicapped child under twenty-one years of age, shall create a substantial risk to the health or safety of the child, by violating a duty of care, protection, or support. It is not a violation of a duty of care, protection, or support under this division when the parent, guardian, custodian, or person having custody or control of a child treats the physical or mental illness or defect of the child by spiritual means through prayer alone, in accordance with the tenets of a recognized religious body.

\textit{OHIO REV. CODE ANN. § 2919.22} (West 2019).
As the court in Spivey noted, a number of factors could explain why the case ended in tragedy. Given the risks of such outcomes, it is unsurprising why prosecutors move forward with child endangerment charges against caregivers who engage in crimes like shoplifting while leaving children unattended in vehicles. Indeed, the defendant in Spivey received an arguably light sentence of 180 days in county jail, largely because the charges were misdemeanor offenses. But in some jurisdictions, prosecutors can prosecute CLUV as a felony as illustrated by Fernandez v. State.

As in Spivey, the defendant in Fernandez shoplifted from a Target while she left a fifteen to eighteen-month-old infant in her care locked in her car. When police rescued the child, “[t]he windows were rolled up; the baby was warm, sweating, and crying; and the child had a soiled diaper.” In addition to CLUV charges, the prosecution sought and obtained a conviction for child abandonment, a felony charge for which the defendant received a sentence of incarceration for two years. The Texas Court of Appeals affirmed both of the convictions.

2. Consequences of CLUV-Related Convictions (n = 9, 14.0%)

Nine cases in the sample dealt with the consequences associated with CLUV convictions. These cases primarily

---

128 Spivey, 2021 WL 3234383, at *3.
129 Id.
130 Whether such prosecutions actually deter others from engaging in such conduct, however, remains an open question.
131 See Spivey, 2021 WL 3234383, at *2.
133 Id.
134 Id.
135 Id. The court suspended the term of incarceration and in lieu of the incarceration, released the defendant to community supervision for five years. Id.
136 See infra notes 209–213 and accompanying text.
questioned the proportionality of the punishment or attempted to expunge of the relevant convictions. Overall, the courts did not take favorably to such claims as evidenced by the fact that appellate courts in all nine of these cases rejected the appellants’ arguments. *State v. Long* is illustrative.138

The defendant in *Long* drove a van for a daycare facility.139 On a day she failed to check that all children were out of the van when she dropped them off in the morning, a twenty-two-month-old girl remained in the van for approximately 7.5 hours until she was found dead in the vehicle at the end of the workday.140 The defendant pled guilty to the low-level felony of reckless homicide in exchange for a two-year sentence as a so-called “standard offender which means that she statutorily could have qualified for the period of incarceration to be suspended while she was on probation.”141 The trial court denied the alternate sentence, citing the seriousness of the offense and the need for other daycare workers to be reminded of what can happen if they abandon their duty to care for the children under their supervision.142 She appealed and the Tennessee Court of Criminal Appeals affirmed, reasoning that:

> Generally, to deny probation or another alternative sentence based on the seriousness of the offense, the offense “as committed, must be ‘especially violent, horrifying, shocking, reprehensible, offensive, or otherwise of an excessive or exaggerated degree,’ and the nature of the offense must outweigh all factors favoring” an alternative sentence.

> We agree with the trial court that the circumstances surrounding this offense are particularly shocking and reprehensible and that the nature of the offense outweighs the factors favoring probation or another alternative sentence. The Defendant pled guilty to reckless homicide for her participation in circumstances leading to the death of . . . a twenty-two-month-old child. The Defendant was charged with the responsibility of picking up young children and delivering them to the daycare center. She clocked out and left on July 21, 1999, while [the

---

139 *Id.*
140 *See id.*
141 *Id.; see also id.* at *3 (“A defendant who ‘is an especially mitigated or standard offender convicted of a Class C, D, or E felony is presumed to be a favorable candidate for alternative sentencing options in the absence of evidence to the contrary.’”) (quoting TENN. CODE ANN. § 40-35-102(6) (2010)).
142 *See id.* at *2.*
child] was still strapped into her car-seat. Due to her tender age, the child was unable to free herself from the car-seat or otherwise remove herself from the van. Her well-being was entirely dependent upon the care and attention of others. Because the Defendant did not fulfill her responsibility of safely delivering [the child] into the daycare center, the child remained in the stifling hot van for seven and one-half hours, where she ultimately died from the heat in the van.143

Cases in which defendants sought to seal their convictions or otherwise expunge their criminal records followed a similar trajectory insofar as courts cited the seriousness of the offense and the need to prioritize the safety of children as grounds for denying such requests—even in cases where the law typically would seal convictions at a certain level of criminal offense.144 Consider what the court said in People v. Nicholas:

The decision of whether or not to seal a record under 160.55(1) is based upon whether or not it is in the interest of justice to do so. There is a body of jurisprudence dealing with leaving a child alone arising from criminal and family court decisions that are helpful in deciding how to apply the “interest of justice” standard in this case where a child was left unattended in a motor vehicle by the defendant.

. . . .

The facts admitted to by Ms. Nichols were that she left an eighteen month old child alone in a vehicle strapped into a car seat on March 16, 2007 at 4:30 p.m. in a public parking lot for over 25 minutes with the motor running while she had a tanning session.

. . . .

The Court finds [from] the conduct of the defendant in this case . . . [that] it was “reasonably foreseeable that extreme harm could come to a young child” left alone at age 18 months in a car with its motor running in a public parking lot at dusk on a March evening over twenty minutes while the defendant attends a tanning session, i.e., sexual predators, carjacking, carbon monoxide, sudden illness of child.

. . . .

The Court, then, finds that the interest of society in the safety and welfare of children by keeping this record unsealed is greater than society’s interest in relieving the defendant of the “stigma” a public arrest record entails.145

143 Id. at *4 (quoting State v. Cleavor, 691 S.W.2d 541, 543 (Tenn. 1985)) (citations omitted).


3. Sufficiency of the Evidence (n = 49, 76.5%)

As previously stated, the majority of cases in the research sample involved appeals challenging the sufficiency of the evidence at trial involving violations of CLUV-specific laws, homicide statutes, or laws criminalizing child endangerment, neglect, or abuse.146

a. Establishing Risk

One of the most common claims in sufficiency of the evidence appeals revolved around questions of whether the actions of the defendants posed significant risks to the well-being of children. Challenges to the application of child endangerment laws on vagueness grounds—because they do “not inform the public that leaving a child in a car unattended constitutes endangering the

welfare of a child”—were universally rejected. Vagueness challenges aside, appellate courts nearly always upheld trial court determinations that leaving children unattended in vehicles presented risk covered by applicable statutes. In fact, many cases involved determinations that children had been exposed to significant risk beyond those attendant to exposure to weather conditions while unattended in vehicles. State v. Obeidi serves as a good example.

In Obeidi, the defendant had left her one and three-year-old children in her SUV. She claimed that she did not expect her children to be in any danger because she planned on being in the store for a short time to buy diapers. Additionally, “she locked the SUV and turned on the car alarm, and she left the windows partially open for air circulation.” However, she was in the store for between twenty and thirty minutes, during which time a witness observed the older child lean out the open window, “pull on the outside door handle, and nearly fall onto the pavement.” Those facts, as well as concerns that the open window could have allowed anyone in the high crime area to have taken the children from the SUV, prompted the witness to call the police who, upon arrival, found the children “fine” just as the defendant returned to her vehicle. The Oregon Court of Appeals affirmed the defendant’s conviction for child neglect under these facts, reasoning that in addition to the risks of the toddler falling out of the open window,

---


148 The court in Watson noted that:

[The jury had competent evidence that defendant left his three and one-half month old baby unattended and lightly clad in an unlocked car for as long as twenty-five minutes. The car was in a shopping center parking lot in the middle of winter and the temperature was cold at the time. The jury also heard defendant’s testimony that he consciously considered as his only options bringing his daughter into the store and carrying her or putting her in a shopping cart without a built-in child seat or leaving her in the car. He stated that he felt she was safest in the car.

From this competent evidence, the jury rationally could have found . . . that defendant consciously disregarded the risks attendant upon leaving the child unattended in a car under these circumstances; and that defendant’s disregard of the risk was a gross deviation from the conduct of a reasonable and prudent person.

149 See Obeidi, 155 P.3d at 81 n.1.

150 See id. at 81.

151 Id.

152 Id.

153 See id. at 81–82.

154 See id.
both children faced risks of assault and abduction because they were left alone in “a busy, high-crime area.”

b. Proving Mens Rea

The other common argument raised in sufficiency of the evidence appeals concerned proof of mens rea. A comparison of three key cases demonstrates how questions of intent typically present in CLUV cases, especially when caregivers claim to have forgotten children.

In both Whitfield v. Commonwealth and State v. Taylor, daycare workers were charged following the deaths of children who had been accidentally left in a daycare vehicle. The defendant in Whitfield was charged and convicted of involuntary manslaughter and felony child neglect stemming from the death of a thirteen-month-old that he left in the daycare van. At trial, he testified in his own defense, admitting that he normally checked the van, but failed to do so on that particular occasion. He also noted:

[The daycare had also trained Whitfield to fill out a logbook in the van to help him keep track of the children he picked up and dropped off at the daycare. Whitfield did not use the van logbook that day, nor had he used it for several months beforehand. Instead, Whitfield admitted, he chose to rely solely on his memory.]

Both charges against Whitfield required proof of criminal negligence, but the way that level of mens rea is defined under Virginia law uses language that conflates traditional elements of recklessness and gross negligence. Indeed, case law

---

155 See id. at 80, 84; see also City of Beachwood v. Hill, No. 93577., 2010 WL 2783140, at *1–4 (Ohio Ct. App. July 15, 2010) (noting risks similar to those in Obeidi regarding CLUV when windows are open and the vehicle is in a high-crime area); cf. Allison v. State, No. 661, 2015, 2016 WL 5462439, at *3 (Del. Sept. 28, 2016) (“In addition to the potential for the children to become dehydrated or overheated [on a hot and humid day], the children were distressed and crying [and one had even vomited]. Leaving young children alone in an unlocked car on an extremely hot day could likely cause physical and mental harm to the children.”).

156 Cases in which courts engaged in statutory interpretation to determine the requisite mens rea for conviction are not included in this section. See infra text accompanying notes 256–262.


159 Whitfield, 702 S.E.2d at 591–92.

160 Id. at 592.

161 Id.

162 See id. at 594. The court cited Noakes v. Commonwealth, 699 S.E.2d 284, 289 (Va. 2010), stating that:

Under Virginia law, criminal negligence occurs ‘when acts of a wanton or willful character, committed or omitted, show a reckless or indifferent disregard of the rights of others, under circumstances reasonably calculated to produce injury, or
paradoxically defines negligence as involving “reckless or indifferent disregard,” but case law allows for this standard to be met through an objective standard showing that “the defendant ‘either knew or should have known the probable results of his/her acts.’” Applying this confusing standard, the court reasoned that the defendant in Whitfield had not experienced a “momentary, inadvertent act of ordinary negligence,” but rather had exhibited a “pattern of reckless indifference” by failing to check the van and not using safeguards like his logbook or the one inside the daycare used to check children in and out.164

By comparison, State v. Taylor involved a miscommunication between two daycare workers. In that case, the defendants left two children unattended in hot cars for roughly two hours and forty minutes when the outdoor temperature was ninety-one degrees Fahrenheit. One child died and the other suffered significant neurological injuries, leading to charges of child abuse resulting in great bodily harm by reckless disregard and child abuse resulting in death by reckless disregard. Both defendants argued that “because they were not aware that the children were left in the vehicle, they could not have consciously disregarded the risk of leaving the children in the car,” which is the traditional standard for recklessness. The Supreme Court of New Mexico agreed with the defendants that a substantial question existed about the sufficiency of the evidence and, therefore, ordered their release during the pendency of their appeal.

In State v. Thompson, a father was convicted of felony child neglect resulting in death after his two-year-old died as a result of being left in the family car for four hours on a day when outside temperatures exceeded eighty degrees Fahrenheit. After a largely sleepless night because the child had a fever and kept his parents up all night, the defendant awoke at 3:00 AM to find his trailer home flooding due to heavy rains. He carried his child to his car and then drove to various locations until the waters

\*which make it not improbable that injury will be occasioned, and the offender knows, or is charged with the knowledge of, the probable result of his or her acts.\*

Id. (internal citations and quotation marks omitted).

163 Id.
164 Id. at 594–95.
165 See Taylor, 491 P.3d at 739.
166 Id.
167 See id. (citing N.M. STAT. ANN. § 30-6-1(D)–(E) (2009)).
168 See id. (citing N.M. STAT. ANN. § 30-6-1(D), (F) (2009)).
169 Id. at 743 (citing MODEL PENAL CODE § 2.02(2)(c) (AM. L. INST. 1985)).
170 Id. at 745. An appeals court had affirmed the defendants' convictions, but the state high court granted certiorari and the case remains pending as of this writing. See generally State v. Taylor, 493 P.3d 463 (N.M. Ct. App. 2021), cert. granted, 504 P.3d 533 (N.M. 2021).
171 State v. Thompson, 647 S.E.2d 526, 528 (W. Va. 2007).
172 Id.
They returned home in the mid-morning. After making it home, the defendant left the boy in the car and went into the trailer to change into dry clothes. He maintained that he left the child “in the car because it was still raining,” the child was still running a fever, and the electricity was out. He claimed that once inside, “he collapsed into sleep unintentionally by reason of physical exhaustion and did not wake up until he heard the restoration of electric power,” at which point the boy had been in the car for four to five hours. The Supreme Court of Appeal of West Virginia affirmed his conviction and his indeterminate sentence of three to fifteen years imprisonment, largely because the mens rea for the offense was lower than recklessness or even gross negligence. Indeed, the jury had been instructed in accordance with a statutory definition of “neglect” that “means the unreasonable failure by a parent, guardian or custodian of a minor child to exercise a minimum degree of care to assure the minor child’s physical safety or health.” Applying what amounts to a negligence standard, the court concluded:

the evidence was there for the jury to conclude that the appellant contributed to the circumstances which led, inexorably, directly to [his son] Luke’s death from hyperthermia. The death was foreseeable. The appellant was aware of his own exhaustion from being up the entire night, and he knew that he was the only adult present to take responsibility for Luke. He could have carried Luke and the car seat into the trailer. Moreover, according to the State, the appellant entered the trailer to wait for [the child’s mother to] return, rather than simply to change clothes.

The *Thomson* case demonstrates that when a caregiver accidentally leaves their child unattended in a vehicle for what might seem perfectly understandable reasons, such actions may result in felony criminal liability, depending on the mens rea requirements in a jurisdiction.

---

173 *Id.*
174 *Id.*
175 *Id.*
176 *Id.* at 528–29.
177 *Id.* at 529, 534. The court relied on statute in noting:

> If any parent, guardian or custodian shall neglect a child under his or her care, custody or control and by such neglect cause the death of said child, then such parent, guardian or custodian shall be guilty of a felony and, upon conviction thereof, shall be fined not less than one thousand dollars nor more than five thousand dollars or committed to the custody of the Division of Corrections for not less than three nor more than fifteen years, or both such fine and imprisonment.

*Id.* (quoting W. VA. CODE § 61-8D-4a(a) (1997)).
178 *Id.* at 533; W. VA. CODE § 61-8D-1(7) (1988).
179 *Id.* at 534.
Children Left Unattended in Vehicles

4. Statutory Construction or Constitutionality (n = 6, 9.4%)

Six cases in the research sample involved questions of the meaning of purportedly ambiguous language in applicable statutes, or challenges to the constitutionality of the laws or the ways in which they were being applied in a specific case.\(^{180}\) Most of these cases involve questions of statutory interpretation, the resolution of which impacted defendants who had been charged under such provisions.

a. What Intent Is Required?

Some state laws clearly identify the level of mens rea that the prosecution must establish for a CLUV-related conviction.\(^{181}\) Other state laws, however, are ambiguous, which has necessitated courts to clarify what level of mens rea, if any, is a required element of such offenses. Consider Michigan’s statute:

A person who is responsible for the care or welfare of a child shall not leave that child unattended in a vehicle for a period of time that poses an unreasonable risk of harm or injury to the child or under circumstances that pose an unreasonable risk of harm or injury to the child.\(^{182}\)

In *People v. Haveman*, Michigan charged the defendant with violating this law.\(^{183}\) The defendant in the case “parked her car in a Walmart parking lot and went inside to shop for one hour, leaving her three and five-year-old children and two dogs inside the vehicle with one window rolled down.”\(^{184}\) An employee saw the children in the car and notified police who, in turn, arrested the mother.\(^{185}\) At her trial for two counts of violating the CLUV law quoted above, she asserted that the statute required proof of specific intent, while the prosecution countered it needed only to prove general intent.\(^{186}\) The court, however, reasoned that because the statute is silent with respect to the requisite mens rea, the CLUV offense carried strict liability. The defendant sought interlocutory review, which the Michigan Court of Appeals

---


\(^{182}\) MICH. COMP. LAWS ANN. § 750.135a(1) (2009).

\(^{183}\) Haveman, 938 N.W.2d at 776.

\(^{184}\) Id.

\(^{185}\) Id.

\(^{186}\) Id.
granted.\textsuperscript{187} After noting that “strict liability is disfavored” for criminal offenses, the court concluded the offense required mens rea, but only of general intent.\textsuperscript{188} The court reasoned that most other child welfare offenses under Michigan law required proof of mens rea.\textsuperscript{189} The court also contended that the punishments—which ranged from up to ninety-three days in jail and up to a $500 fine if a child sustained no injuries, to between ten to fifteen years in prison and $5,000 to $10,000 in fines if a child was seriously injured or died—also weighed in favor of interpreting the statutory intent as requiring proof of mens rea.\textsuperscript{190}

b. The Effect of Potentially Overlapping Statutes

Recall that \textit{Shouse v. Commonwealth} illustrated how a defendant’s use of drugs or alcohol might cause prosecutors to view a CLUV incident as warranting moving forward with charges.\textsuperscript{191} With that decision having been made, prosecutors secured three convictions in \textit{Shouse}—one for wanton murder, one for second-degree criminal abuse, and another for first-degree wanton endangerment.\textsuperscript{192} One of the key questions on appeal concerned the propriety of the wanton murder conviction.\textsuperscript{193}

Kentucky does not have a specific statute targeting CLUV, generally. But the state did amend its second-degree manslaughter statute to include a provision aimed at deaths which occur as a result of CLUV:

(1) A person is guilty of manslaughter in the second degree when he wantonly causes the death of another person, including but not limited to situations where the death results from the person’s:

(a) Operation of a motor vehicle;

\begin{itemize}
\item \textsuperscript{187} \textit{Id}.
\item \textsuperscript{188} \textit{See id.} at 779, 782. The court noted that the CLUV law could not be properly considered a public welfare offense since that type of crime curtails conduct that runs “contrary to the interest of public safety,” such as those “dangers caused by the industrial revolution, increased traffic, the congestion of cities, and the wide distribution of goods.” \textit{Id.} at 781 (internal citations omitted).
\item \textsuperscript{189} \textit{See id.} at 779–81.
\item \textsuperscript{190} \textit{See id.} at 781–82; \textit{see also} People v. Maynor, 683 N.W.2d 565, 569 (Mich. 2004) (holding that when CLUV results in the death of a child and the defendant is charged with first degree child abuse, the prosecution must prove “not only that defendant intended to [commit the act], but also that . . . defendant intended to cause serious physical harm or . . . knew that serious physical harm would be caused [by the act]”); \textit{see also id.} (“A person is guilty of child abuse in the first degree if the person knowingly or intentionally causes serious physical harm or serious mental harm to a child. Child abuse in the first degree is a felony punishable by imprisonment for not more than 15 years.”) (quoting Mich. Comp. LAWS ANN. § 750.136b(2) (West 2022)).
\item \textsuperscript{191} \textit{See Shouse v. Commonwealth}, 481 S.W.3d 480, 482–83 (Ky. 2015); \textit{see also supra notes} 110–115 and accompanying text.
\item \textsuperscript{192} \textit{Shouse}, 481 S.W.3d at 482–83.
\item \textsuperscript{193} \textit{Id.} at 483.
\end{itemize}
(b) Leaving a child under the age of eight (8) years in a motor vehicle under circumstances which manifest an extreme indifference to human life and which create a grave risk of death to the child, thereby causing the death of the child.\textsuperscript{194}

As the statutory language makes clear, wanton conduct is an element of the offense. However, that same element is part and parcel of wanton murder under Kentucky law, which provides that a person commits murder under circumstances when he, she, or they “wantonly engages in conduct which creates a grave risk of death to another person and thereby causes the death of another person.”\textsuperscript{195} The latter provision, which parallels the common law depraved heart murder rule, creates liability for an unintentional death that results from grossly reckless conduct—the conscious disregard of a known risk of death or serious injury to another person under circumstances manifesting an extreme indifference to the value of human life.\textsuperscript{196} In \textit{Shouse}, prosecutors sought and obtained a conviction for wanton murder, a more serious offense than second-degree manslaughter.\textsuperscript{197} On appeal, the Supreme Court of Kentucky set aside the wanton murder charge, reasoning that the language included in the amendments to the second-degree manslaughter statute evidenced legislative intent that an unintentional death resulting from leaving a child age eight or younger in a vehicle, “under circumstances which manifest an extreme indifference to human life and which create a grave risk of death to the child,” constitutes manslaughter in the second degree, not murder.\textsuperscript{198}

The court in \textit{Shouse} also vacated the defendant’s wanton endangerment charge because that offense requires conduct “manifesting extreme indifference to the value of human life.”\textsuperscript{199} The court reasoned as follows:

There is a clear distinction between driving, even under the influence of drugs and in a vehicle with a spare donut tire on the car, and leaving a child abandoned in a car overnight to die. And no harm came from her driving at that point, so it is difficult to say that there was a substantial danger of death or serious physical injury. Otherwise, driving with a

\textsuperscript{194} See \textit{Ky. Rev. Stat. Ann.} § 507.040(1) (2000). The court in \textit{Shouse} quoted from treatise to emphasize that the legislature had amended the second-degree manslaughter statute “to provide explicitly for homicide coverage of the situation where a person leaves a child under 8 years of age in a motor vehicle and in so doing causes its death.” \textit{Shouse}, 481 S.W.3d at 483 (quoting \textit{Robert G. Lawson & William H. Fortune, Kentucky Criminal Law} § 8.4(a), at 31 (Supp. 2006)).


\textsuperscript{197} \textit{Shouse}, 481 S.W.3d at 482–83.

\textsuperscript{198} \textit{Id.} at 482–88.

\textsuperscript{199} \textit{Id.} at 489.
donut tire replacing a flat to get home from a dinner where one had consumed a glass of wine would per se be first-degree wanton endangerment. Certainly a possibility of injury existed, but further proof of the degree of danger is necessary for the higher offense.200

In State v. Ducker, the defendant was charged with two counts of first degree murder, stemming from the deaths of her thirteen-month-old and twenty-three-month-old children.201 She contended that she did not intend to kill her children, but the state’s murder statute allowed for a first-degree murder conviction for the unintentional death of a child resulting from aggravated child abuse or aggravated child neglect.202 A jury acquitted her on the first-degree murder charges, but convicted her of two counts of aggravated child abuse.203 On appeal, she attacked the latter conviction on the grounds that it could not be a lesser included offense of murder, primarily because aggravated child abuse requires a higher level of mens rea than the recklessness or gross recklessness for an unintentional homicide.204 The court rejected her argument and concluded that a lesser included offense analysis was misplaced because the state also recognizes that “a lesser grade or class of the charged offense[,]” even with different elements, statutorily satisfies the requirements of lesser included offenses:205

Tennessee law recognizes two types of lesser offenses that may be included in the offense charged in an indictment and, may, therefore, form the basis for a conviction: a lesser grade or class of the charged offense and a lesser included offense. The two, though similar, are not synonymous. An offense is “lesser included” in another “if the elements

---

200 Id. (emphasis omitted).
203 Ducker, 1999 WL 160981, at *14. Child abuse, neglect, or endangerment becomes aggravated under Tennessee law when, among other factors, the underlying conduct results in serious bodily injury to a child. TENN. CODE ANN. § 39-15-402(a)(1) (1989). The non-aggravated forms of child abuse and neglect require knowledge as their requisite levels of mens rea:
(a) Any person who knowingly, other than by accidental means, treats a child under eighteen (18) years of age in such a manner as to inflict injury commits a Class A misdemeanor; provided, however, that, if the abused child is eight (8) years of age or less, the penalty is a Class D felony.
(b) Any person who knowingly abuses or neglects a child under eighteen (18) years of age, so as to adversely affect the child’s health and welfare, commits a Class A misdemeanor; provided, that, if the abused or neglected child is eight (8) years of age or less, the penalty is a Class E felony.
(c)(1) A parent or custodian of a child eight (8) years of age or less commits child endangerment who knowingly exposes such child to or knowingly fails to protect such child from abuse or neglect resulting in physical injury or imminent danger to the child.
204 Ducker, 1999 WL 160981, at *14.
205 See id.
of the greater offense, as those elements are set forth in the indictment, include but are not congruent with, all the elements of the lesser.” However, a lesser “grade or class” of offense is established by the legislature and is determined simply by reading the statutory provisions.

Necessarily included within the offense of aggravated child abuse . . . is the offense of child abuse and neglect. . . . [The murder statute] provides that, if fairly raised by the evidence, child abuse is a lesser offense of first degree murder. This provision, despite the appellant’s assertions, is necessarily encompassed within the aggravated child abuse statute. Accordingly, we hold that, in the present case, aggravated child abuse is a lesser included offense of first degree murder. . . .

As the summaries of Shouse and Ducker should make clear, the cases centered on questions of different levels of potential liability in the wake of a CLUV incident causing the death of a child. Other cases involved similar questions with regard to the level of liability that should attach for child endangerment, child abuse, or child neglect when a child is rescued before conditions in a vehicle lead to the child’s death. Recall, for instance, State v. Fernandez, a case in which the defendant shoplifted while she left an infant unattended in her locked car. In addition to the CLUV charges discussed previously, the Texas Court of Appeals also affirmed a conviction on the defendant’s child abandonment charges over the defendant’s challenge that the felony count of child abandonment and the misdemeanor CLUV charge were in pari materia (on the same subject). The court quoted both statutes:

The child abandonment statute provides, “A person commits an offense if, having custody, care, or control of a child younger than 15 years, he intentionally abandons the child in any place under circumstances that expose the child to an unreasonable risk of harm.” The other statute at issue in this case provides, “A person commits an offense if he intentionally or knowingly leaves a child in a motor vehicle for longer than five minutes, knowing that the child is (1) younger than seven years of age; and (2) not attended by an individual in the vehicle who is 14 years of age or older.”

The court determined that the two statutes were not in pari materia because the felony child abandonment charge could only be committed “by a person having custody, care, or control of a child,” whereas the misdemeanor CLUV offense could be “committed by anyone.” Moreover, the two offenses differ with respect to essential elements, with the felony requiring proof of

---

206 Id. (citations omitted) (quoting State v. Trusty, 919 S.W.2d 305, 310 (Tenn. 1996)).
208 See supra notes 132–136 and accompanying text.
209 See Fernandez, 269 S.W.3d at 66–67 (first quoting TEX. PENAL CODE ANN. § 22.041(b) (1993) (child abandonment); and then TEX. PENAL CODE ANN. § 22.10(a) (1984) (CLUV)).
210 Id. at 66 (citations omitted).
211 Id. at 67.
exposure to some risk of unreasonable harm, whereas the CLUV offense does not.\textsuperscript{212} Thus, the two statutes need not be considered together and, therefore, separate convictions were appropriate.\textsuperscript{213}

c. Impermisssible Statutory Presumptions

One case in the research sample raised an interesting issue about the limits of CLUV statutes' evidentiary presumptions. In \textit{People v. Jordan}, the defendant had left his daughter in his car in sub-freezing temperatures while he went to go buy college textbooks.\textsuperscript{214} A security officer at the university heard the baby crying and proceeded to contact the proper emergency authorities.\textsuperscript{215} Jordan was ultimately convicted under the Illinois child endangerment statute, which contains a specific provision relating to leaving a child unattended in a vehicle for ten minutes or longer.\textsuperscript{216} The key issue in the case concerned the wording of a subsection of the statute, which provided for a "rebuttable presumption that a person committed the offense if he or she left a child 6 years of age or younger unattended in a motor vehicle for more than 10 minutes."\textsuperscript{217} The defendant argued this provision was unconstitutional because it operated to create a mandatory presumption of guilt.\textsuperscript{218} The U.S. Supreme Court previously held that conclusive presumptions and mandatory rebuttal presumptions that shift the burden of persuasion to the defendant are unconstitutional because they contravene the presumption of innocence.\textsuperscript{219} Additionally, Illinois case law had established the same holding with respect to mandatory rebuttable presumptions that shift the burden of production to the defendant.\textsuperscript{220} In light of these facts, the court in \textit{Jordan} declared that portion of the statute unconstitutional.\textsuperscript{221} Had that subsection of the statute been

\footnotesize
\begin{itemize}
  \item \textsuperscript{212} See id.
  \item \textsuperscript{213} See id. at 68.
  \item \textsuperscript{214} People v. Jordan, 843 N.E.2d 870, 873–74 (Ill. 2006).
  \item \textsuperscript{215} \textit{Id.} at 873.
  \item \textsuperscript{216} \textit{Id.} at 872–73 (quoting 720 ILL. COMP. STAT. ANN. 5/12-21.6 (West 2002)). For the text of the statute, see \textit{supra} note 120.
  \item \textsuperscript{217} \textit{Jordan}, 843 N.E.2d at 872–73 (quoting 720 ILL. COMP. STAT. ANN. 5/12-21.6(b) (West 2002)).
  \item \textsuperscript{218} See id. at 876.
  \item \textsuperscript{219} See id. at 876–77 (citing Sandstrom v. Montana, 442 U.S. 510, 521–23 (1979)).
  \item \textsuperscript{220} \textit{Id.} at 877 (citing People v. Watts, 692 N.E.2d 315, 323 (Ill. 1998)). The court in \textit{Watts} reasoned:

A production-shifting presumption places a burden on the defendant to come forward with a certain quantum of evidence to overcome the presumption. If the defendant does not satisfy that burden, the judge is required, in effect, to direct a verdict against the defendant on the element which is proved by the use of the presumption. This result conflicts with the longstanding rule that a verdict may not be constitutionally directed against a defendant in a criminal case.

\textit{Watts}, 692 N.E.2d at 323.
  \item \textsuperscript{221} \textit{Jordan}, 843 N.E.2d at 879.
\end{itemize}
permitted to stand, unless the defense submitted evidence to the contrary, it could have had the effect of requiring the trier-of-fact to accept that a child had been left in a vehicle unattended for ten minutes or more. The court severed that unconstitutional provision and concluded that even in the absence of any presumption, there was still sufficient evidence to support a conviction for child endangerment:

[A] rational trier of fact could have found that defendant knowingly endangered his infant daughter's life or health by leaving her unattended in his vehicle. Several factors bear upon that determination including the setting where the vehicle was parked, the weather conditions [a windy day when the outside air temperature was in the twenties], and the amount of time defendant left his daughter alone in the vehicle [as long as forty minutes].222

As a result of this finding, the court remanded the case for retrial—a result that is permissible since retrial of defendants after a finding of sufficient evidence of guilt does not violate double jeopardy.223

IV. DISCUSSION

We consider one of our null findings to be among the most interesting results in our study, namely that the circumstances under which a CLUV incident occurred was not a statistically significant predictor of case outcome. Put differently, on the whole, courts did not treat cases where a child was forgotten in a vehicle any differently than they did when a child was knowingly left unattended in a vehicle. Given the lack of difference in case outcomes across these two factual predicate situations, it seems imperative to raise more awareness about the dangers of CLUV. We summarize several ways to do that without resorting to the use of criminal law in Part IV.B.224 In this subsection, we discuss the significant finding from our quantitative and qualitative analyses.

A. Effectiveness of Formal Social Control Via Criminal Laws

Overall, it is difficult to ascertain the effectiveness of laws to reduce the incidence of CLUV because, as the analyses of the cases in the research sample demonstrate, the overwhelming majority of cases in which cases are prosecuted involve more extreme

---

222 Id. at 877, 879. For other cases assessing risk, see supra notes 147–155. The court also noted that the child faced other risks beyond those associated with the weather:

[I]t is an unfortunate fact of modern urban life that the more populated the area, the greater the likelihood that some ill will befall a young child who is left unattended in a public place. A young child unattended in a public setting is easy prey for social predators who may happen by. Jordan, 843 N.E.2d at 879.

223 Id. at 881 (citing, inter alia, People v. Placek, 704 N.E.2d 393, 396–97 (Ill. 1998)).

224 See infra notes 238–262 and accompanying text.
situations, typically those in which children either died or sustained serious injuries. By contrast, cases in which charges were filed for CLUV when children were timely recused were far less represented in the research sample. So, it may be that laws targeting CLUV are having some effect. After all, there is no way to gauge whether fewer children are actually being left in vehicles either knowingly or accidentally if the children survive without sustaining injuries. Certainly, such incidents would not be reflected in databases like those maintained by No Heat Stroke, which only tracks pediatric vehicular heatstroke cases. Still, the qualitative data analyses in the present study suggest that some laws are operating as presumably intended and others require some modifications.

*Shouse v. Commonwealth* is one of the cases that perhaps best illustrates the potential efficacy of a well-implemented CLUV law, even though it only applies to child fatalities rather than a broad range of CLUV situations. The judicial analysis of the Kentucky law at issue in *Shouse* suggests that legislators realized traditional avenues for prosecuting child deaths are often overly punitive, especially toward parents grieving the unintentional deaths of their children. As interpreted by the Kentucky Supreme Court, the state CLUV law provides a mechanism to impose criminal liability for caregivers’ recklessness while also applying less severe punishment than would be applicable for other types of unintentional homicides resulting from grossly reckless conduct. This was done perhaps out of a recognition that caregivers, especially those who are parents, likely experience intense guilt and continuous grief when children die as a result of their own behavior.

By contrast to the way the CLUV homicide law in Kentucky operates in the wake of the *Shouse* decision, it appears that the CLUV provisions in the laws of some other states could be amended for clarity. The results also suggest that improvements are needed with regard to the way such cases are handled. Two notable findings about final case outcomes on appeal support these conclusions.

First, there were statistically significant differences in the prevailing parties across the types of issues being considered in each case. Perhaps due to the sympathy for child victims and moral outrage for caregivers whose conduct caused injuries or death to

---

225 See supra notes 34–46 and accompanying text.
226 See *Shouse v. Commonwealth*, 481 S.W.3d 480, 489 (Ky. 2015).
227 *Id.* at 484.
those victims, courts in the research sample unanimously affirmed lower-court sentences, including those in which sentencing judges (1) opted against diversion programs for which defendants were eligible, (2) imposed incarceration terms rather than probation when the latter was available, and (3) imposed longer terms of incarceration than those in presumptive sentencing ranges. Additionally, judges denied requests in all cases where defendants sought to seal or expunge CLUV-related convictions.

However, the fact that defendants were successful in twice as many cases as prosecutors when raising issues of statutory construction supports the idea of rewriting some states’ CLUV laws with more precision. Such CLUV laws can be made more precise by either clarifying the meaning of mens rea or better delineating how CLUV laws work in conjunction with other criminal laws—without violating double jeopardy and associated principles.

Second, the success rate for appeals on the sufficiency of the evidence claims suggests that the processing of CLUV cases may need improvement. Even though the majority of cases in which the appellant raised a sufficiency of the evidence claim ultimately resulted in an outcome favoring the prosecution, the 32.6% reversal rate for these claims in the research sample is quadruple the national reversal rate of 8.1% for all other crimes. This finding

---

229 See, e.g., Tom Googhegan, *Hot Car Deaths: The Children Left Behind*, BBC News (July 22, 2014), http://www.bbc.com/news/magazine-28214266 [http://perma.cc/8WV6-QKTW] (reporting that “[t]he response to these kinds of cases is commonly vitriolic” and quoting a father whose actions in a CLUV incident caused his seventeen-month-old child’s death as saying that people “want[ed] to crucify [him] for what [he] did and that [he] was one of those people before it happened to [him]”); see also Gouveia, *supra* note 7; Washabaugh, *supra* note 7, at 201. Such emotions might also explain why the age of the children left unattended in vehicles was significantly related to case outcome. Very young, helpless children are particularly sympathetic victims. CLUV incidents involving them, therefore, might impact triers-of-fact during verdict deliberations, as well as appellate judges when reviewing both the sufficiency of the evidence and the propriety of sentences.


236 See *Waters et al.*, *supra* note 105, at 6 fig.3.
reinforces the notion that CLUV cases can be complex, especially when caregivers clearly did not intend or desire any harm to their children. But it also suggests that prosecutors may move forward with cases where the facts do not necessarily support the charges. Whether due to outrage over unnecessary deaths of children, moral condemnation about caregivers’ use of drugs or alcohol or engagement in criminal activities, the desire to use CLUV cases as examples for general deterrence purposes, or a combination of these reasons and, perhaps, other reasons not mentioned, the high reversal rate of CLUV-related convictions on appeal suggests either an overcharging or undervaluation of case facts.237

B. Alternative Social Controls

Volumes of research strongly indicate that informal social controls are often more effective mechanisms for changing behavior than are formal social controls.238 Indeed, “[e]mpirical evidence on the deterrent effects of punishment remains speculative and inconclusive, and the ability of formal punishment alone to deter crime appears to be quite limited.”239 By contrast:

- informal social controls have proven to be effective in curtailing juvenile delinquency;
- reducing illicit drug cultivation, distribution, and use;
- reducing alcohol abuse;
- reducing domestic violence;
- reducing the incidence of driving under the influence;
- increasing worker productivity;
- and even helping to reduce recidivism among sex offenders.240

Given that most CLUV-related injuries or deaths occur accidentally, there is a strong argument that formal social controls cannot deter the underlying behaviors—especially when caregivers forget children in vehicles. But informal social controls can play an important role in raising awareness about the risks of knowingly leaving children unattended in vehicles. Additionally, other styles of social control and technology hold promise for raising awareness to reduce the incidence of forgetting children in vehicles.

1. Public Service Campaigns

In a recent study of parental attitudes concerning pediatric vehicular heatstroke, nearly one out of four parents reported having left their children unattended in vehicles, many of whom were unaware of the associated dangers. This finding suggests that we need more media and other public service campaigns to teach people about the risks of CLUV so that caregivers might better understand those risks—even for brief periods of time or on days when they perceive the weather to be mild.

Such educational ventures could take several forms. Law enforcement agencies and schools could advertise about CLUV dangers. Mass media—including television, radio, magazines, billboards, and social media platforms—could warn about the dangers of CLUV in public service announcements (like the ones created by Kids and Car Safety and the National Highway Traffic Safety Administration’s “Where’s Baby? Look Before You Lock” campaign) in much of the same ways they have done for decades to raise awareness about the dangers of driving while impaired. Similarly, driver education programs could include modules on CLUV, just as they do for driving while impaired.

Such campaigns should be consistent with Williams and

---


244 Numerous studies report that such campaigns reduce alcohol-impaired driving. E.g., John P. Murry, Jr., Antonie Stam & John L. Lastovicka, Evaluating an Anti-Drinking and Driving Advertising Campaign with a Sample Survey and Time Series Intervention Analysis, 88 J. AM. STAT. ASSN 50, 50, 55–56 (1993); Kimberly P. Whittam, William O. Dwyer, Patricia W. Simpson & Frank C. Leeming, Effectiveness of a Media Campaign to Reduce Traffic Crashes Involving Young Drivers, 36 J. APPLIED SOC. PSYCH. 614, 615–16, 625 (2006). Also, at least one study found that increased volumes of drinking and driving public service announcements were associated with statistically significant decreases in alcohol-related traffic fatalities. See Jeff Niederdeppe, Rosemary Avery & Emily N. Miller, Alcohol-Control Public Service Announcements (PSAs) and Drunk-Driving Fatal Accidents in the United States, 1996–2010, 99 PREVENTIVE MED. 320, 320–24 (2017).

Grundstein’s recommendations to target common misperceptions and reduce caretakers’ cognitive dissonance.246

2. Technology

Technology can help prevent CLUV incidents.247 Indeed, some car manufacturers have already integrated technologies that use sensors to detect children (or pets, for that matter) in vehicles and then alert drivers to their presence.248 Some of these devices are quite simple, while others are high-tech:

Vayyar’s 3D imaging sensor can detect movements throughout the car after the engine is off. If there is someone detected, the system will alert the registered driver via text message or phone call, sound the car alarm or, if the car is electric, activate the air conditioner. Other technologies can detect carbon dioxide, weight, vitals, temperature, and more.249

Technology that alerts drivers about potential CLUV incidents is one of the few solutions that the federal government could mandate. The Hot Cars Act seeks to do just that by requiring all new vehicles to include technology that would alert a driver to a passenger in the backseat of a vehicle.250 But even if enacted, it would take years to make a notable difference due to the fact that older cars would need to be phased out of use and be replaced by newer vehicles equipped with such technology. Moreover, such technology would only address unintentional CLUV. It would not impact the incidence of caregivers knowingly leaving children unattended in vehicles. And, of course, technology failures occur all the time, rendering reliance on technological solutions less than ideal. Collectively, these shortcomings underscore the importance of the educational and public service efforts previously summarized.251

249 Id.
250 See id.; see also Hot Cars Act of 2021, H.R. 3164, 117th Cong. (2021).
251 We note that one or two automakers have implemented technology that may reduce the likelihood of harm associated with leaving a child (or pet) in a vehicle. For
3. Other Styles of Formal Social Control

There will undoubtedly be people who feel that penal social controls are necessary to combat the problems associated with CLUV. We certainly agree that this solution would be appropriate in certain cases, such as when children sustain injuries or die. Indeed, we even support penal social controls for repeat offenders in CLUV incidents in which children were not harmed at all. But for first-time offenders whose children are not injured, we believe the compensatory style of social control holds promise in curbing repeated behavior through a combination of civil sanctions and educational interventions. Specifically, states could enact laws for first-time offenders causing no harm to children that are modeled after California’s Kaitlyn’s Law, which is contained in the California Vehicle Code—not the penal code—and is sanctioned as a civil infraction.252 Notably, that statute provides that caretakers found in violation of the law may be required to attend an educational program addressing the risks of leaving a child unattended in a vehicle.253

Critics have argued that laws like California’s Kaitlyn’s Law amount to little more than legal showmanship because they typically result in the imposition of minor fines or even only verbal warnings.254 One possible remedy to address such concerns would be to mirror Michigan’s approach, which increases the sanction for subsequent violations.255 We caution, however, against two features of Michigan’s law.

example, Tesla implemented a feature known as “Dog Mode” which keeps the cabin of the vehicle climate-controlled while displaying the current temperature inside the car along with a message on the primary on-board monitor saying that the owner will be back soon. See Brian Wang, Tesla Adds Sentry and Dog Modes, NEXT BIG FUTURE (Feb. 20, 2019), http://www.nextbigfuture.com/2019/02/tesla-adds-sentry-and-dog-modes.html [http://perma.cc/V65G-L5WL]. Ford has filed patents for similar technology, but has yet to implement it. See Justin Banner, Ford Barks Up Tesla’s Tree, Considering Dog Mode–Style Pet Feature, MOTOR TREND (Feb. 17, 2022), http://www.motortrend.com/news/ford-patent-dog-pet-mode-feature/ [http://perma.cc/9R67-662B]. Neither of these technologies were designed to protect humans, although they might be adapted for such use. We see this as potentially problematic because it signals that leaving a child unattended in a vehicle is acceptable under certain circumstances when, in fact, it is never safe to do so.  

252 CAL. VEH. CODE § 15620 (West 2022).
253 Id.
255 Michigan’s approach is as follows:
   (2) A person who violates this section is guilty of a crime as follows:
   (a) Except as otherwise provided in subdivisions (b) to (d), the person is guilty of a misdemeanor punishable by imprisonment for not more than 93 days or a fine of not more than $500.00, or both.
   (b) If the violation results in physical harm other than serious physical harm to the child, the person is guilty of a misdemeanor punishable by
First, as previously discussed, Michigan’s law is less than clear about the mens rea required and the attendant circumstance element regarding what might constitute an “unreasonable risk of harm or injury.” For that matter, California’s Kaitlyn’s Law also uses indeterminate terminology about risk. The results of our qualitative analysis lead us to believe that more clarity is needed than such generalizations provide. To avoid uncertain questions of fact about those circumstances that pose a risk to a child left unattended in a vehicle, a certain length of time is preferable, such as the lengths of time used in Florida and Texas. Given that serious injuries can occur after just five minutes in a hot vehicle, we recommend that a specific length of time serve as the trigger for liability.

Second, Michigan classifies a first offense of CLUV as a misdemeanor punishable by potential jail time and a fine. We think that would be appropriate for a second offense, with subsequent offenses increasing in the lengths of potential incarceration and in the amounts of fines. But given the widespread ignorance about the dangers of CLUV even for short periods of time or on mild days, a first offense could be dealt with as a civil infraction, rather than a misdemeanor, just as California law provides. Moreover, this approach has the added benefit of allowing for the imposition of strict liability for a first offense, thereby obviating any disputes about mens rea.

imprisonment for not more than 1 year or a fine of not more than 1,000.00, or both.
(c) If the violation results in serious physical harm to the child, the person is guilty of a felony punishable by imprisonment for not more than 10 years or a fine of not more than 5,000.00, or both.
(d) If the violation results in the death of the child, the person is guilty of a felony punishable by imprisonment for not more than 15 years or a fine of not more than 10,000.00, or both.

256 See id. § 750.135a(1); supra notes 182–190 and accompanying text.
257 Cal. Veh. Code § 15620 (West 2022) (prohibiting CLUV “[w]here there are conditions that present a significant risk to the child’s health or safety”).
259 See Impact of Dangerous Microclimate Conditions, supra note 33, at 105.
261 See supra notes 23–33 and accompanying text.
262 California law provides the following:

A violation of [this law] is an infraction punishable by a fine of . . . one hundred dollars ($100), except that the court may reduce or waive the fine if the defendant establishes to the satisfaction of the court that he or she is economically disadvantaged and the court, instead, refers the defendant to a community education program that includes education on the dangers of leaving young children unattended in motor vehicles, and provides certification of completion of that program.

while not engendering any concerns about overcriminalization in the absence of criminal intent.

C. A Suggested Model Statute

Based on our reading of all of the cases in the research sample, we offer the following model statute that addresses the full range of CLUV outcomes—from minor incidents in which no child is hurt all the way through unintentional deaths caused by knowingly leaving a child unattended in a vehicle. Our proposal grades the offense and associated penalties for repeat offenders of minor incidents in which no child is harmed. When a child is harmed as a result of knowing or reckless conduct, our proposed statute also grades the offense and associated penalties based on the degree of harm the child encounters. Finally, to avoid potential double jeopardy issues, the model statute makes it clear that intentional harms may be prosecuted as other offenses, but unintentional harms should not.

(1) A parent, legal guardian, or other person responsible for a child who is eight (8) years of age or younger may not leave that child inside a motor vehicle unattended or unsupervised by a person who is fourteen (14) years of age or older and not legally incapacitated, under any of the following circumstances:

(a) For a period of five (5) minutes or longer;
(b) For any period of time if the vehicle’s engine is running or the vehicle’s keys are in the vehicle, or both; or
(c) Where there are any conditions that present a significant risk to the child’s health or safety as assessed from the viewpoint of the ordinary, reasonable, prudent person.

(2) Provided that no physical harm has come to the child as a result of being left in the vehicle under any of the circumstances specified in paragraph (1), any person who violates the provisions of paragraph (1) for the first time is guilty of a noncriminal, strict liability infraction. All such first-time violators shall be punished by a fine of $250.00, except that the court may reduce or waive the fine if the defendant establishes to the satisfaction of the court that he or she is economically disadvantaged. In addition, regardless of whether or not a fine is imposed, all first-time violators shall be mandated to participate in an education program on the dangers of leaving young children unattended in motor vehicles.

(3) Except as provided in paragraph (2), any person who, acting with purpose, knowledge, recklessness, or criminal negligence, violates the provisions of paragraph (1) without causing any physical harm to a child is guilty of a crime as follows:

(a) A second violation of the provisions of paragraph (1) constitutes a misdemeanor punishable by a term of probation of up to one (1) year and/or a fine of not more than $500.00, or both. Additionally,
the court may, at its discretion, suspended the driver’s license of any
defendant convicted under this section for up to ninety (90) days.

(b) A third violation of the provisions of paragraph (1) constitutes
a misdemeanor punishable by a term of incarceration of up to
ninety (90) days in jail, a fine of not more than $1,000.00, or both.
In addition, the court shall order the suspension of the driver’s
license of any defendant convicted under this section for a period
of one hundred eighty (180) days.

(4) Except as provided in paragraphs (2) or (3), any person who, acting
with knowledge or recklessness, violates the provisions of paragraph (1)
is guilty of a crime as follows:

(a) Except as otherwise provided in subdivisions (b) to (d), the
person is guilty of a misdemeanor punishable by incarceration of
up to one hundred eighty (180) days in jail, a fine of not more than
$2,500.00, or both. In addition, the court shall order the suspension
of the driver’s license of any defendant convicted under this section
for a period of one (1) year.

(b) If the violation results in physical harm other than serious
physical harm to the child, the person is guilty of a misdemeanor
punishable by incarceration of up to one (1) year in jail, a fine of
not more than $3,500.00, or both. In addition, the court shall order
the suspension of the driver’s license of any defendant charged
under this section for a period of one (1) year.

(c) If the violation results in serious, albeit unintentional, physical
harm to the child that causes permanent disability or permanent
disfigurement, the person is guilty of a felony punishable by
imprisonment for not more than two (2) years, a fine of not more
than $5,000.00, or both.

(d) If the violation results in the unintentional death of the child, the
person is guilty of a felony punishable by imprisonment for not more
than five (5) years, a fine of not more than $10,000.00, or both.

(5) Although unintentional harms to a child caused by violation of the
provisions of paragraph (1) are to be prosecuted under this subdivision
of the criminal code, nothing in this subdivision precludes prosecution
for any applicable criminal charges stemming from acts or omissions
specifically intended to cause physical injury to or the death of a child
left in a vehicle in violation of the provisions of paragraph (1).

CONCLUSION

Responding to incidents in which a child is left unattended
in a vehicle poses difficult challenges for the criminal legal
system, especially given the prevalence of children accidentally
being forgotten in vehicles. However, even when caregivers
knowingly make the decision to leave children in vehicles, they
typically underestimate the effects that solar radiation can have
on the climate inside a vehicle, as well as the length of time
needed for children to become distressed. To effectively reduce
CLUV-related injuries and deaths, we should employ a range of
strategies, including: (1) CLUV-specific civil violations for first-
time offenders whose actions cause no harm; (2) CLUV-specific
criminal penalties for repeat offenders and those whose actions
unintentionally harm children; (3) educational and public service
campaigns to raise awareness of the risks associated with CLUV,
even for a short period of time on temperate days; and (4)
technological mitigations to prevent drivers from accidentally
leaving children in vehicles.
In the Name of the Environment Part III: 
CEQA, Housing, and the Rule of Law

Jennifer Hernandez*

This is the third study of all state court lawsuits filed under the California Environmental Quality Act (“CEQA”); this Study examined lawsuits filed statewide over three years, between 2019 and 2021. All three studies identified housing as the top target of CEQA lawsuits challenging agency approvals of private projects. California’s housing crisis has caused the state to have the worst housing-adjusted poverty rate in the United States; California also continues to have the highest rate, and highest number, of unsheltered homeless residents. Housing production has remained essentially flat (at about 110,000 housing units per year) notwithstanding the enactment of more than one hundred new housing laws since 2017; the state still needs about three million more homes. Although CEQA’s status quo defenders assert that CEQA is not a material factor in housing production, this Study confirms that, in 2020 alone, CEQA lawsuits sought to block approximately 48,000 approved housing units statewide—just under half of the state’s total housing production. Many housing laws also mandated that local and regional agencies adopt and implement plans to accommodate more housing. CEQA lawsuits filed during the study period challenged agency housing plans that allowed more than one million new housing units. Non-housing projects to accommodate housing and population growth, such as transportation and water infrastructure, are also a major target of CEQA lawsuits. CEQA lawsuits (and lawsuit claims) relating to

* Jennifer Hernandez practices environmental and land use law in the California offices of Holland & Knight. Many other members of Holland & Knight’s West Coast Land Use and Environment and Real Estate Practice Groups contributed to the study of CEQA lawsuits evaluated in this report, including Nicholas Quinlivan, Scott Levin Gesundheit, Nathaniel R. Bernstein, Deborah Brundy, Melanie Chaewsky, Emily Warfield, John H. Irons, Stevens A. Carey, Ariel B. Robinson, Brad Brownlow, Daniel Golub, David Friedman, Norman Carlin, Brian Bunker, and Emily Lieban. While the author is grateful to these contributors and to other parties who are focused on the need to modernize CEQA, the opinions and recommendations in this Article are the author’s and should not be attributed to any other person or organization, to Holland & Knight, or any client of the firm. This Article cites to media reports and other specified sources for factual information about examples of CEQA lawsuits and the litigation practices of individuals and groups—they were not independently investigated by the authors.
climate change, including greenhouse gas emissions ("GHG") and vehicle miles travelled, a top topic of CEQA lawsuits, even though California already has the lowest per capita GHG in the nation and has enacted scores of GHG and climate change laws and regulations. The study includes data, and examples, of all CEQA lawsuits filed during the study period, to explain how CEQA works today—not historically, and not rhetorically.

The study also examines how the unpredictability of CEQA lawsuit outcomes has created a de facto, low-cost, no-risk strategy for project opponents to preserve the environmental status quo and block even benign and beneficial projects until litigation (inclusive of appeals) is completed—typically about in four to five years. This judicial outcome uncertainty has made lenders, investors, and grantors unwilling to fund projects while CEQA lawsuits remain pending, thereby allowing CEQA petitioners to avoid the judicial preliminary injunction process, in which they must persuade a judge that they are likely to prevail on the merits, and will suffer irreparable harm unless the project is halted. A judge can also require petitioners to post a bond to cover delay damages if their lawsuit is ultimately determined to be meritless. Judicial uncertainty in CEQA lawsuits has, in practice, meant that judges can only stand by for the eighteen to twenty-four months of delay that petitioners obtain by the simple act of a filing a lawsuit and paying a small court filing fee.

CEQA lawsuit outcome uncertainty is also a profoundly influential factor in how much time and money is spent on CEQA compliance (especially for projects more likely to be sued, such as housing in wealthier communities, as was shown in the second of this CEQA study series). The study examines CEQA jurisprudence in contrast to the administrative jurisprudential factors typically applied to statutes and regulations, and explains the practical consequences of judicially-imposed expansions of CEQA—and judicially-rejected enacted legislation imposing, by statute, interpretive and remedy constraints on judicial outcomes in CEQA lawsuits. One potential explanation for this judicial rejection of the plain language of statutes, such as prohibiting courts from imposing a CEQA remedy to stop construction of a legislative office building in Sacramento unless the office building caused health or safety harms or adversely affected a previously-unknown significant tribal resource; notwithstanding this statutory language, the appellate court stopped this construction project based on historic resource and aesthetic concerns.
This Article notes that CEQA lawsuits are also filed as “writs”—not ordinary civil lawsuits—which have a long history and tradition of vesting extraordinary discretion in the judiciary, which acts as a separate and co-equal branch of government and has an independent role in enforcing the Rule of Law, including through use of its equitable authorities. Courts enforce statutes all the time, however, in both writ and non-writ proceedings, and the key attributes of the Rule of Law—including knowing in advance what the law requires—have not constrained many judicial decisions that expand CEQA well beyond what is required by any clear, discernable compliance mandate in CEQA statutes or implementing regulations.

Legislative reform of CEQA, unless acceptable to powerful special interests such as certain labor and environmental organizations, remains mired in Sacramento’s politics. The actual pattern of CEQA lawsuits, reflected in this and the prior two studies, should give pause to CEQA’s status quo defenders, who—like this author—often personally profit from CEQA’s unbounded costs and schedules. CEQA’s most visible status quo defenders assert their allegiance to the environment and “environmental justice” (though not other civil rights); they have been buoyed by special interests who wield CEQA as a sword to protect proprietary (and often economic) interests.

CEQA’s statutory bias is to preserve the status quo, even when the status quo is causing ongoing harms to people (including hard working families who never voted to abandon the California Dream of homeownership but have been priced out by the housing crisis), or the environment (which needs change to prevent forest fires and catastrophic floods, and achieve massive change to energy production and climate adaptation). With multi-year studies followed by an over four-year litigation slog, CEQA’s foundational prioritization of procedural perfection undermines solving urgent housing, civil rights and environmental priorities.

California has enacted thousands of environmental laws and regulations since CEQA was signed into law in 1970. CEQA’s extended adolescent fixation on process over progress—inclusive of unpredictable, grandiose, and chaos-inducing behaviors and outcomes—needs to grow up. This Study makes the same three CEQA reform suggestions as prior studies, and adds one more. First, end anonymous CEQA lawsuits: parties filing CEQA lawsuits need to identify who they are, and show that they are suing to protect the environment, just like they’ve always had to do when suing under federal environmental lawsuits. Second, end
duplicative CEQA lawsuits: once a project or plan has completed the CEQA process, no new CEQA lawsuits can be filed as the project is constructed and plan is implemented—progress must occur and process must end. Third, match the remedy to the crime: if an agency made a mistake and didn’t study an impact enough, then the appropriate judicial remedy in CEQA—as already prescribed in the CEQA statute itself—is for a judge to require more study and mitigation, without rescinding project approvals and requiring agencies and applicants to re-do the CEQA process for another two years, followed by another six years of litigation after that. Housing delayed in housing denied, and a deficient traffic study shouldn’t result in a six year re-run of CEQA processing. Fourth, and new for this Study: this author’s plea for the judiciary to return to the norms of administrative law jurisprudence, and cannons of statutory construction, when deciding CEQA cases. Simply: no Legislative reform will be effective without judicial outcome predictability consistent with the Rule of Law.

INTRODUCTION ...........................................................................62
I. STUDY BACKGROUND AND METHODOLOGY .....................70
II. CEQA v. HOUSING ...............................................................75
   A. CEQA v. Apartments ....................................................79
   B. CEQA v. Agency Plans and Ordinances Allowing More Housing ....................................................81
   C. CEQA v. Homeownerships for Middle Income Families ....................................................88
   D. CEQA v. Students .........................................................94
   E. CEQA v. Old People .......................................................96
   F. CEQA v. Homelessness ................................................96
   G. CEQA v. Single Family Homes/Casitas .....................97
III. CEQA v. EVERYTHING ELSE (NON-HOUSING) ..............100
      1. CEQA v. Water Equity .............................................102
      2. CEQA v. Streets and Sidewalks ...............................109
      3. CEQA v. Schools ..................................................110
      4. CEQA v. Parks/Trails .............................................111
B. CEQA v. Non-Residential Private Projects .......... 111
   1. CEQA v. Cannabis ............................................. 112
   2. CEQA v. Warehouse/E-Commerce ......................... 113
   3. CEQA v. Renewable Energy ................................... 115
   4. CEQA v. Agriculture ........................................... 116
   5. CEQA v. Fun ....................................................... 116
   6. CEQA v. Retail .................................................... 117
   7. CEQA v. Hotels ................................................... 117
   8. CEQA v. Not Much Else ...................................... 117

C. CEQA v. Non-Residential Agency Plans and
   Regulations .............................................................. 118
   1. CEQA v. Climate Change ................................ 119

IV. CEQA AND THE RULE OF LAW ............................................ 122
   A. Administrative Law Jurisprudence v. CEQA Jurisprudence .................................................. 124
      1. Student Noise and Conventional Administrative Law Practice ........................................ 125
         a. Plain Language Rule ....................................... 126
         b. Deference to Expert Administrative Agency Interpretation .......................................... 126
         c. Deference to Lead Agency Analytical Methodology and Factual Findings ...................... 126
         d. Consistency with Other Statutes, and with Constitutional Protections ........................ 127
      2. CEQA Jurisprudential Deviations from Administrative Law Norms: Undergraduate Student Dorm Occupancy Example ........................................ 127
      3. Next Steps with Social Noise and CEQA .............. 128

B. CEQA and The Rule of Law ........................................... 130
   1. Off the Rails: Leading Court Cases Creating Modern CEQA Jurisprudence ......................... 132
   2. 1993 and Beyond: Legislature’s Largely Failed Attempts to Restore Administrative Law Jurisprudence to CEQA .................................................. 141
INTRODUCTION

This is the third in a series of how California’s venerable environmental law, the California Environmental Quality Act (“CEQA”), enacted in 1970, is actually litigated in the real world. All three studies examined all CEQA lawsuits filed statewide, and each concluded that the most frequent target of CEQA lawsuits was housing approved in existing communities. The studies spanned 2010-2012, 2013-2015, and this current study period of 2019-2021.

In our second study, published in the Hastings Environmental Law Journal, I observed that CEQA lawsuits “provide a uniquely powerful legal tool to block, delay, or leverage economic and other agendas,” and “is now the tool of choice for resisting change that would accommodate more people in existing communities.” These observations, and other data and observations from our second study, were quoted at length in a recent First District Court of Appeal case involving a twenty-five-year odyssey and 900-page Environmental Impact Report (“EIR”) for a thirty-four single family home project on a parcel in Marin County adjacent to the wealthy town of Tiburon (median home price, $2,862,177). As the Court of Appeal observed, “all of these . . . observations are vindicated in this woeful record before us.”

California has the highest poverty rate, and highest homeless population, in the nation. There’s a common reason for these shocking humanitarian failures by the fourth largest economy on

---

3 Tiburon Open Space Comm. v. County of Marin, 294 Cal. Rptr. 3d 56, 122 (Ct. App. 2022).
the planet, long governed by a supermajority of Democrat state officers and Legislators, in one of the deepest green states in the nation: state policies block housing that’s affordable to its residents, with leaders and advocates defending state policies in the name of the environment (and now climate), even when they expressly acknowledge the exclusionary harms their policy choices inflict on younger families, communities of color, and middle income (including union) workers.

In December 2022, the California Air Resources Board (“CARB”)—California’s leading air quality and climate agency—adopted a “Scoping Plan” that included scores of policy choices and mandates to reduce greenhouse gas (“GHG”) emissions using a metric that primarily counts electricity and petroleum consumption by California’s residents and businesses. For example, the Scoping Plan counts GHG from cement and other building products produced in California and does not count GHG from imported cement and other products. The Scoping Plan includes hundreds of GHG and climate policy choices, and was unanimously-adopted by a Board consisting entirely of appointees of Democratic party leaders in the state.

CARB’s policy choices, as the Scoping Plan expressly concludes, is that households making $100,000 or less will bear a disproportionately high cost burden to pay for the state’s climate policy choices (including housing). CARB further acknowledges that these middle and lower income households are far more likely to be comprised of Black or Latino residents than White or Asian residents. I describe the disparate race-and-class-based harms inflicted on Californians under the climate change environmental

---

7 CAL. AIR RES. BD., 2022 SCOPING PLAN FOR ACHIEVING CARBON NEUTRALITY 158 (2022).
8 Id. at 86, 208.
11 See id. at 126–27.
banner, in which recently-invented CEQA climate change “impacts” of people who occupy new housing play a pivotal role, in *Green Jim Crow*. In *Green Jim Crow* lays bare the racist attributes of core climate policies that are implemented through CEQA, such as the elevation of Vehicle Miles Travelled (“VMT”) as an environmental “impact” even when the vehicle being driven is a zero emission electric car. The goal of the VMT policy was to focus more housing near high frequency transit, and increase housing density, to reduce GHG emissions from the ordinary activities of Californians. In actuality, the neighborhoods slated for redevelopment into high density housing under the VMT policies largely overlapped with majority Black and ethnic minority neighborhoods first mapped by federal mortgage insurance “redlining” maps to deny residents of these neighborhoods access to the attainable homeownership programs offered to White residents (and veterans) under the New Deal and beyond. The VMT incentive policies increase gentrification and displacement, and incentivize exceptionally high cost housing ($1,000 or more per square foot, resulting in $3,500 or more monthly rents or over $1 million or more condos), to the direct detriment of displaced communities of color and other median/low income households. As discussed further, *infra*, VMT “mitigation” obligations add costs of $50,000 or more to new housing in non-transit locations, which are most frequently used to subsidize public transit or construct bike lanes for other people, somewhere else, even if not proximate to the new housing. Like the cost of land, labor, and building materials, mitigation costs imposed by agencies increase the cost of producing new housing, but VMT mitigation costs are most often assessed against families forced to “drive until they qualify” for housing they can afford to buy or rent, who must still drive to get to work (like more than ninety-five percent of Californians). People buying or renting these lower cost suburban homes are most likely to be the median/middle income households (now majority minority) who cannot afford high density housing in the fraction of one percent of California located within a half mile from high frequency bus stops, rail stations, or ferry terminals.

---

13 See id.
14 See id.
15 See id.
16 See generally id.
17 See generally id.
This Part Three of our *In the Name of the Environment* series provides further evidence of California’s anti-housing environmental/climate agenda: as described in more detail in Part II (CEQA v. Housing), just 7 of the 514 lawsuits in this Study’s dataset sought to block 1,079,347 planned housing units (half to one-third of California’s estimated housing shortfall). Lawsuits filed in just one year (2020) sought to block just under 48,000 approved housing units (the equivalent of just under half of California’s total annual housing production). The entrenched strength of these environmental/climate anti-housing stakeholders is all the more remarkable given the Governor’s conclusion that the state has 3.5 million fewer housing units than it needs,18 and given the scores of new laws enacted by the Legislature and signed by the current and former governor to spur increased housing production.

CEQA remains a revered cornerstone of California’s environmental laws, even as all three *In the Name of the Environment* CEQA studies confirm that CEQA lawsuits are most often aimed at blocking housing and climate priorities purportedly supported by the state’s elected leaders. More academic researchers, including once-ardent CEQA status quo defenders, have independently confirmed the accuracy of the data in our studies—and increasingly have also acknowledged its use as an anti-housing exclusionary tool by wealthier communities. For example, as explained by UC Berkeley Law Professor Eric Biber in *CEQA and Socioeconomic Impacts: Why Expanding CEQA to Cover Socioeconomic Impacts Might Harm Equity Goals*, in a comment criticizing an appellate court decision to roll back undergraduate enrollment at UC Berkeley:

> In research I have helped work on about how CEQA and local land-use law is implemented for housing projects in California, we have found evidence that litigation and administrative appeals are more common in wealthier neighborhoods fighting projects. This suggests it is more likely that more privileged communities will use socioeconomic impact analysis challenges under CEQA to stop needed housing projects, housing that is needed to resolve the state’s dire housing crisis.19

Professor Biber’s research observation mirrors my own. In our second study, *California Environmental Quality Act Lawsuits and*

---


California’s Housing Crisis, we mapped the location of the nearly 14,000 housing units challenged in a swath of Southern California to show that anti-housing CEQA lawsuits are indeed far more common in Whiter, wealthier, and healthier neighborhoods.\textsuperscript{20}

Challenges to the CEQA status quo have also become more frequent, and less politically verboten, by elected leaders. For example, CEQA has been described by a prominent State Senator and pro-housing production leader as “the law that swallowed California.”\textsuperscript{21} This Study provides direct evidentiary support for the accuracy of that observation in Parts II (CEQA and Housing) and Part III (CEQA and Everything Else).

Part IV examines CEQA judicial precedents, including the unwillingness of many courts to apply longstanding administrative jurisprudential canons such as deferring to the plain language of the CEQA statute, in anti-housing and other CEQA lawsuits. CEQA jurisprudence—reported appellate and Supreme Court decisions—has made the outcome of CEQA lawsuits entirely unpredictable, as we first reported in an earlier study examining fifteen years of judicial outcomes in CEQA lawsuits.\textsuperscript{22}

No statute should be so ambiguous, uncertain, or incomprehensible that our institutions and people don’t know what’s even required under our Rule of Law system, as described by the American Bar Association.\textsuperscript{23} Unless agencies—and those regulated by agencies including project applicants—know what the law requires, the adequacy of CEQA compliance more closely resembles judicial outcomes of core Constitutional disputes such as the fuzzy line between free speech and obscenity, which prompted U.S. Supreme Court Justice Potter Stewart’s famous test for what is obscene: “[he] know it when [he] sees it.”\textsuperscript{24}

\textsuperscript{20} Hernandez, In the Name of the Environment II: 2013-2015, supra note 1, at 30–32.
For years after our CEQA judicial outcome study was published, and consistently through today, CEQA status quo defenders excuse unpredictable judicial outcomes and persist in blaming agencies for losing CEQA lawsuits because they are “not doing CEQA the right way.”25 However, two prominent law school professors (and longtime CEQA status quo defenders) opined in a recent amicus brief defending a longtime CEQA petitioner lawyer.26 The petitioner lawyer’s conduct met the legal test of “malicious prosecution” when she intentionally filed a meritless CEQA lawsuit to block a single family home project in San Anselmo (near Tiburon, also in Marin County).27 In the brief, the law school professors explained that the “fact” of CEQA litigation uncertainty is “universally acknowledged,” citing to the accuracy of data we gathered in an earlier CEQA study showing roughly 50/50 odds of a project opponent beating an agency in a CEQA lawsuit.28

Well into fifteen years of our comprehensive study of CEQA, it is impossible to explain CEQA litigation patterns without highlighting the role the judiciary has and continues to play in expansively and creatively applying CEQA to identify new analytical and other requirements that are not expressly written into CEQA, the CEQA Guidelines, or any prior published case comprising CEQA jurisprudence.

This third In the Name of the Environment study recommends, in Part V, that the California Supreme Court grant review of a recent UC Berkeley appellate court decision, which elevates for the first time the “social noise” of partying undergraduates in unbuilt dorms as a CEQA impact, and take the opportunity created by this review to revisit CEQA jurisprudence.

It has been fifty-one years since the California Supreme Court’s first CEQA decision, Friends of Mammoth, involves a condo project near Mammoth ski resort.29 In that case, the Court


26 See Brief for Richard M. Frank & Sean B. Hecht as Amici Curiae Supporting Appellants, Jenkins v. Brandt-Hawley, 302 Cal. Rptr. 3d 883 (Ct. App. 2022) (No. A162852) [hereinafter Frank & Hecht Amicus Brief].

27 Id.

28 Id.

directed lower courts to interpret CEQA broadly to protect the environment. That decision is the most often cited by courts that decline to apply the plain language of CEQA's statutes. *Friends of Mammoth* continues to serve as the rationale for courts' willingness to require the most draconian of CEQA remedies—rescission of project approval pending more CEQA legislation. This draconian remedy has been applied even to housing that has been constructed, and occupied, while the lawsuit was pending. Justice Chin, in a dissent to an anti-housing CEQA lawsuit in which all justices agreed that CEQA was not a population control statute, stated: “We have caution[ed] that rules regulating the protection of the environment must not be subverted into an instrument for the oppression and delay of social, economic, or recreational development and advancement.”

The outsize use of CEQA lawsuits to block housing in existing communities is just one cause of the housing crisis, and the Legislature is dutifully enacting dozens of laws each year to try to spur housing production to address the state's multi-million housing unit shortfall. This is not California's first CEQA versus housing battle of the "super-statutes," as another law professor has quipped. The new housing legislation is again at risk of being crushed by CEQA, repeating the last round where the legislation directed more and faster housing approvals in the 1980s, only to be crushed by CEQA judicial decisions which included judicial rejection of plain language statutory directives beginning with the wholesale rejection of most 1993 legislative reforms to CEQA and continuing through the UC Berkeley

---

30 See generally *Friends of Mammoth*, 502 P.2d at 1049.
31 Ctr. for Biological Diversity v. Dep't of Fish & Wildlife, 361 P.3d 342, 367 (2015) (Chin, J., dissenting) (quoting Citizens of Goleta Valley v. Bd. of Supervisors, 801 P.2d 1161, 1175 (1990)). The majority agreed that “CEQA is not intended as a population control measure.” Id. at 350.
Unless the California Supreme Court and appellate courts revisit their own “business-as-usual” CEQA jurisprudence, pro-housing, pro-civil rights, and pro-climate resiliency statutory mandates and funding priorities will be first delayed and then crushed by status quo defenders in CEQA lawsuits. Fortunately, the judiciary’s pathway to success is less politically fraught than the Legislature’s navigation through the swarm of special interests that use CEQA to advance their own economic and (non-environmental) policy goals.

In what I fervently hope will be the last in our *In the Name of the Environment* series, I write this to give voice to struggling, hard-working Californians who do not have a swarm of special interest lobbyists but do need and deserve to be able to work hard and buy a home (and shouldn’t need to visit their kids and grandkids via videoconference because the kids couldn’t afford to stay in California). I write this to give voice to lower income Californians scrambling for too few affordable units who endure decade-long lottery delays for taxpayer-funded affordable housing (and do not want to rent a one bedroom cottage in someone else’s backyard to raise their family). I write this for residents who need (and pay among the highest taxes in the country to use) effective, reliable, and affordable water, transportation, and energy infrastructure, as well as public services like parks, schools, and public safety. For the people who could not just de-camp to Hawaii during the state’s extended COVID lock-down, for those without fancy college degrees who make a living by showing up and doing a job, not just tapping on a keyboard. For these people—my families and the tens of millions of families like mine—we urgently need the state’s elected leaders and our distinguished judiciary to please restore CEQA to ordinary administrative law jurisprudence, and allow critically-needed housing, climate resilient infrastructure, water supplies, and public services to be built in full compliance with the thousands of environmental protection statutes and regulations adopted since 1970—and stop allowing CEQA to be the massive Not in My Back Yard (“NIMBY”) status quo defender (and special interest extortion tool) that it has evolved into over the past fifty-two years.

---

29 See Make UC A Good Neighbor v. Regents of Univ. of Cal., 304 Cal. Rptr. 3d 834 (Ct. App. 2023).
I. STUDY BACKGROUND AND METHODOLOGY

CEQA requires that any party who files a lawsuit alleging noncompliance with CEQA must send a copy of that lawsuit to the California Attorney General (“Cal AG”). For all data gathered in all three of our In the Name of the Environment studies, we sent Public Records Act requests to the Cal AG asking for copies of all CEQA lawsuits filed during each of our three, three-year study periods (2010–2012, 2013–2015, and 2019–2021). Each petition was then reviewed, with pertinent data (such as project location, type of agency action/project challenged, etc.) entered into datasheets, and then compiled into the categories reported in each study. The third data set comprising this Study is reported in Parts III and IV below.

For those unfamiliar with CEQA, we start with a very brief summary of this law and suggest that readers review Getting Started with CEQA, published by the Governor’s Office of

---

36 Tom Meyer, CEQA Cartoon.
37 See CAL. PUB. RES. CODE § 21167.7 (West 2023) (requiring parties to furnish a copy of the petition, as well as any amended or supplemental petition, to the Attorney General’s Office within ten days after filing the pleading).
38 As with past years, copies of each CEQA lawsuit, and datasheets, can be made available for in-person review in the author’s law office by appointment.
Planning & Research ("OPR"). OPR is the state agency assigned by CEQA to develop "guidelines"—which serve as the equivalent for most purposes of regulations—to provide more detailed and practical directions on CEQA compliance requirements.

CEQA is both a procedural statute, requiring analysis and public disclosure of the environmental consequences of proposed agency actions, and a substantive statute, requiring that public agencies fully consider public comments as well as avoid or minimize to the greatest extent feasible significant adverse environmental impact. Only after an agency requires all such "feasible" means of avoiding or reducing an environmental impact is an agency nevertheless allowed to approve a project based on an "overriding" social (e.g., affordable housing), economic (e.g., job-creating), legal (e.g., emergency response to a fire or broken bridge), or technological (e.g., energy conserving LED lighting retrofits), benefits of a project.

CEQA applies to both public agency approvals of their own plans, regulations, and policies and to public agency decisions to approve or fund projects undertaken by the private parties from homeowners to large corporations. Scores of statutory exemptions to CEQA have been approved over the past decades, which are typically limited to politically favored projects, and also include numerous eligibility criteria and restrictions which render many "unicorns"—much discussed, rarely if ever seen in practice. There are also limited regulatory (or "categorical") exemptions from CEQA for project categories that "normally" would not result in adverse environmental impacts. Whether a project qualifies for either a statutory or regulatory exemption can also be challenged in a CEQA lawsuit.

A CEQA lawsuit challenges whether an agency has properly complied with CEQA, but in most cases is intended to—and does—block construction of the approved agency or private party project.

---

40 See id.
until the lawsuit is resolved. There are two reasons for this practical consequence. First, as recently acknowledged in an amicus filing by two longtime CEQA practitioners (both of whom were tenured law school professors at the University of California—one of whom after decades of senior service in the California Attorney General’s office and the other having since moved on to the Earthjustice environmental advocacy organization), the outcome of judicial decisions in CEQA cases is entirely unpredictable. Second, the most common judicial remedy in CEQA decisions against agencies (which happen in about fifty percent of CEQA appellate court cases), is that the project approval is rescinded pending further CEQA processing and re-approval. Even completed apartment projects with tenants in occupancy have been ordered vacated and have been left vacant for years based on a CEQA deficiency identified by a judge or appellate court, often for aesthetic or other non-polluting and non-safety reasons, years after the project was approved.

CEQA lawsuits typically require two to five or more years to resolve, with one lawsuit involving a single family home project on an existing lot in Berkeley that was unanimously supported by adjacent neighbors, the appointed Planning Commission, and the elected City Council, tied up in the courts for eleven years. The Berkeley family homeowner won the lawsuit, but raised their

---

46 Frank & Hecht Amicus Brief, supra note 28, at 12–13.
children elsewhere as their dream house remained mired in litigation and has to this date never been built.\(^{51}\)

In our prior two studies, *In the Name of the Environment: Litigation Abuse Under CEQA* (the “2010–2012 Study”)\(^{52}\) and *California Environmental Quality Act Lawsuits and California’s Housing Crisis* (the “2013–2015 Study”),\(^ {53}\) we showed how CEQA lawsuits have become weaponized to block environmentally beneficial as well as benign projects. For example, in our 2010-2012 Study we showed that CEQA lawsuits are rarely (thirteen percent) filed by recognized environmental advocacy organizations such as the Sierra Club or Center for Biological Diversity,\(^ {54}\) and are instead almost always filed by either individuals, or new and often informal organizations (e.g., named, “Save Fifth Street”) with no identified funding source which was created for the purpose of opposing the challenged project.\(^ {55}\) Through investigative journalists and concurrent media reports, we were able to show that these shadowy new organizations were fronts for anonymous neighbors (e.g., a Berkeley homeowner who opposed the remodel of the community library), competitors (e.g., warring gas station owners), and labor (e.g., retail clerk and construction unions).\(^ {56}\) Unlike all other similar state and federal environmental laws, CEQA lawsuits can be filed by anonymous entities whose primary litigation objective is not protecting the environment.\(^ {57}\)

Our prior two studies also showed that, in stark contrast to the old growth forest clear cuts, major chemical factories, and massive freeways under consideration when CEQA was adopted in 1970, modern CEQA lawsuits are primarily filed to block housing and public infrastructure projects in existing neighborhoods, especially cities.\(^ {58}\) Increased traffic congestion, construction noise, changes to the “character of a community,” and


\(^{52}\) Hernandez, *In the Name of the Environment I: 2010-2012*, supra note 22.


\(^{55}\) See id.

\(^{56}\) See id. at 19, 93.

\(^{57}\) See id. at 24.

other unremarkable characteristics of a growing population, thriving job market, and vibrant but evolving community largely replaced the suite of “environmental” impacts at issue when CEQA was adopted, such as causing the extinction of a endangered species, spewing vast quantities of pollution into the air or water, or destroying the scenic vista of a national park.59

As CEQA reached middle-age (forty years old in 2010), as shown in the first two of our In the Name of the Environment series, CEQA lawsuits were far more likely to be used in “micro-environment” neighborhood disputes, like challenging the renovation of an elementary school cafeteria, the installation of all-weather turf on a public park soccer field, remodels of single family homes, the construction of apartments in existing neighborhoods, and even the addition of neighborhood-conforming single family homes on infill sites like former elementary schools and golf courses. As one commenter has noted in The Atlantic, CEQA has evolved into “a system that subjects even humdrum infill proposals to obtuse multi-binder reports and shady dealings, leaving a housing-affordability crisis in its wake.”60

CEQA lawsuits only challenge approved projects: CEQA follows investment capital.61 In the 2010–2012 Study, for example, we showed that during the state’s first major federal infusion of public funding for transit and renewable energy during the Obama Administration, more CEQA lawsuits challenged solar projects than natural gas power, industrial, and mining projects combined—and more CEQA lawsuits challenged public transit than public highway projects.62

But blocking housing remains CEQA’s most fecund litigation practice. California’s housing crisis has been an acute problem for decades.63 In both earlier studies, we showed that housing was the top target of CEQA lawsuits challenging private

61 See Hernandez, In the Name of the Environment I: 2010-2012, supra note 22, at 43; cf. id. at 22 (“It should come as no surprise that banks making construction loans, and government agencies making time-sensitive grant and appropriations decisions, usually decline to fund projects while a CEQA lawsuit is pending.”).
Public agency plans and zoning decisions that allowed new housing—as well as public agency school and water infrastructure needed for new homes—were the top target of CEQA lawsuits challenging public agency projects. In a deep-dive into anti-housing CEQA lawsuits in the state’s most populous Southern California region, in our 2018 Study we also showed that CEQA lawsuits against housing were more likely to be filed against apartments located near public transit in the region’s wealthier, whiter, and healthier neighborhoods. As has now been widely recognized, CEQA is the most formidable legal obstacle to restoring an adequate housing supply to California.

This third installment in our In the Name of the Environment series again affirms the ongoing pattern of CEQA lawsuits filed against environmental benign and environmentally beneficial housing, renewable energy, and climate resilient infrastructure projects. We first recap the 2023 Study findings, as part of this ongoing pattern, in Part I.

II. CEQA V. HOUSING

In our preview of Anti-Housing CEQA Lawsuits filed in 2020, we tabulated all approved housing units that were challenged in CEQA lawsuits: nearly 48,000 individual housing units were challenged, which, in turn, is nearly half of all of the housing produced statewide in 2020.

An even more startling anti-housing CEQA lawsuit statistic that emerged from a review of all three years in our dataset are lawsuits targeting regional and local agency plans to allow more housing; both the amount of housing needed, and the timing and content of these agency plans, are prescribed by state housing laws. For example, in a 2008 climate bill (SB 375), the regional agencies responsible for managing transportation improvements were charged with developing “Sustainable Communities Strategies” (“SCS”) for coordinating land use and transportation,

---

64 See Hernandez, In the Name of the Environment I: 2010-2012, supra note 22, at 10, 12–15; see also Hernandez, In the Name of the Environment II: 2013-2015, supra note 1, at 29–34.

65 See id. at 26 fig.1. Public services and infrastructure account for 15% of CEQA lawsuits, agency plan/regulation accounts for 19%, and water accounts for 3%. Id. These three categories total 37% of all CEQA lawsuits. Id. Only some of these three categories allow for new housing, totaling about 27% of this 37%. Id. Housing challenges were 29%. Id.


while providing for planned residential and economic growth, all while achieving the state’s aggressive climate change and GHG reduction targets. Notoriously NIMBY advocates targeted the 9-county Bay Area SCS, which accommodated 441,176 new housing units. A separate group targeted the Sacramento SCS in a CEQA lawsuit; that SCS included 133,512 new housing units. Using CEQA to target climate-friendly new housing is a solid example of how CEQA is no longer in alignment with current state environmental priorities.

Cities and counties are also required to adopt “housing elements” in their General Plans, and make other conforming General Plan changes, to accommodate the housing assigned to that jurisdiction under state housing laws. This 2019-2021 study tabulates these anti-housing lawsuits targeting state-mandated General Plan Housing Element updates, which ranged from the massive (City of Los Angeles, assigned 456,643 housing units), to the miniscule (anti-housing opponents in Del Rey Oaks in Monterey County objected to adding 86 new housing units). These lawsuits challenged housing in mid-size cities (Moreno Valley, 11,627 units), and in rural and mountain counties where housing demand surged with COVID refugees from high-cost tiny apartments in San Francisco, such as Calaveras (1,096 units), Placer (7,854 units), and El Dorado (5,353 units).

Anti-housing CEQA lawsuits challenging just two regional SCS climate plans, and six county and city general plans collectively sought to block 1,079,347 new housing units—and nearly a third of Governor Newsom’s inaugural proclamation of a 3.5 million housing shortfall.

To put this in perspective, it is important to recognize that “housing delayed is housing denied” for those unable to find

---

72 See Newsom, supra note 18.
housing they can afford for themselves and their families, and that meets their household’s needs for good schools, accessible jobs, and homeownership. California’s failure to build about three million new homes over the past decades has caused the state to have the highest poverty rate in the nation (inclusive of housing costs), according to the U.S. Census Bureau.\textsuperscript{74}

The housing crisis inflicts the most direct harm on younger Californians, and communities of color, who struggle to pay routine monthly costs even with median and above-median incomes.\textsuperscript{75} Even older homeowners experience the housing crisis, as children and grandchildren move to states with more balanced housing markets.

Essential workers like teachers, healthcare, construction, and retail workers depart for states that still offer attainable homeownership for middle-income households.\textsuperscript{76} The housing crisis has also infected every segment of the California economy, with high housing costs cited as the leading reason why hundreds of thousands of people (net of in-migration) have moved out of California in recent years—a historic reversal of the state’s longtime population growth pattern.\textsuperscript{77}

Our comprehensive three-year evaluation reveals that tabulating actual approved housing units severely undercounts the magnitude of anti-housing CEQA lawsuits. Of the 512 CEQA lawsuits filed during the study period, Figure 1 shows that 198 challenged housing. Of these, 164 challenged housing in cities or on University of California campuses. Only 34 challenged housing on unincorporated county lands outside city boundaries; these

\textsuperscript{74} Morgan Keith, California has the Highest Poverty Level of all States in the US, According to US Census Bureau Data, BUS. INSIDER (Sept. 14, 2021), http://www.businessinsider.com/california-has-highest-poverty-level-in-the-us-census-bureau-2021-9 [http://perma.cc/6T7M-BHKX].


county projects include multi-family housing projects located in higher density unincorporated county neighborhoods (including neighborhoods interspersed between incorporated cities which are served by public transit).

**Figure 1: Locations Targeted in 198 Anti-Housing CEQA Lawsuits**

- Cities (74%)
- College Campuses (9%)
- Unincorporated County Land (17%)

**Figure 2: 198 Anti-Housing CEQA Lawsuits**

- Agency Plans/Regs (21%)
- Homeless (3%)
- Master Plan Community (6%)
- Multi-Family (31%)
- Neighborhood Community (16%)
- Senior Housing (2%)
- Single Family Home/ADU (8%)
- Student Housing (9%)
- Townhomes/Small Subdivisions (4%)
Figure 2 includes all categories of housing approvals challenged under CEQA, including student dormitories, residential facilities restricted to seniors, housing for those experiencing homelessness, income-restricted housing affordable to low and very low-income households, and housing for everyone else. It also includes regional and local agency housing approvals mandated by state laws, notably:

- State climate law\(^{78}\) requiring regions to adopt Sustainable Communities Strategies to reduce GHG contributing to global climate change;
- Regional Housing Needs Allocation (“RHNA”)\(^{79}\) laws requiring regions, cities, and counties to adopt land use plans and zoning ordinances to accommodate a state-assigned allocation of new housing units every eight years; and
- Affirmatively Furthering Fair Housing (“AFFH”),\(^{80}\) requiring that new housing be dispersed throughout a community, including in wealthier neighborhoods that have traditionally been regulated to allow only more costly single-family homes.

A. CEQA v. Apartments

Figure 2 shows that the top target of all CEQA lawsuits filed against housing during the study period are multi-family apartment projects, and Figure 3 shows that anti-housing CEQA lawsuits are the largest category of all lawsuits filed under CEQA. This apartment category includes all housing projects that are restricted to low-income residents, as well as “mixed income” apartments that reserve a percentage of units for lower income households, and “mixed use” projects in urban markets that may include retail or office space on the ground floor. All of these challenged apartment projects are located within urbanized neighborhoods of cities and almost all displace an existing use like a parking lot or single-story commercial or retail building.

As we showed in our 2013-2015 Study,\textsuperscript{81} which dove deeply into CEQA lawsuits where most Californians live (the Los Angeles region), most (78\%) of these challenged multi-family projects were located in wealthier, Whiter neighborhoods, and many (70\%) are proximate to existing or planned public transit (bus or train/BART/metro stops). As explained in greater detail in Green Jim Crow,\textsuperscript{82} land costs, agency-imposed fees, and exactions are generally higher in these urbanized areas. High housing costs are inherent with high cost urban land in desirable locations that displaces existing uses.\textsuperscript{83} The building costs to construct “mid-rise” multi-family projects of up to six stories and “high-rise” apartments over six stories are three to seven times higher due to far more costly structural and operational components, such as elevators and utility systems under various building, earthquake safety, emergency, accessibility, energy, and other mandated components.\textsuperscript{84} These transit rich neighborhoods are also located in cities with outsized job centers such as downtowns and are more likely to impose higher housing fees and more costly exactions.\textsuperscript{85} Apartment projects tend to be most economically feasible in higher wage, higher amenity (e.g., restaurants and other retail services), and neighborhoods with high housing prices and severe housing supply shortfalls for both low income and market rate residents. These apartment projects are also more likely to be in wealthier communities with incumbent homeowners or businesses with an interest in protecting “their” environmental status quo, and the resources to file CEQA lawsuits. Due to high insurance rates, and demanding mortgage rules, almost all multi-family projects are built as rental apartments instead of for-sale condominiums.\textsuperscript{86} Even “affordable” apartments built for low income families in these locations, with this multi-family building typology, cost in excess of $1 million each to produce.\textsuperscript{87} Monthly rental costs for non-subsidized households top $4,000, often without parking, and are “affordable” using the standard benchmark of spending no

\begin{itemize}
  \item \textsuperscript{81} See generally Hernandez, In the Name of the Environment II: 2013-2015, supra note 1.
  \item \textsuperscript{82} See Hernandez, Green Jim Crow supra note 12.
  \item \textsuperscript{83} Id.
  \item \textsuperscript{84} See id. at n.14.
  \item \textsuperscript{85} See Keith, supra note 74.
  \item \textsuperscript{87} Liam Dillon & Ben Poston, Affordable Housing in California Now Routinely Tops $1 Million Per Apartment to Build, L.A. TIMES (June 20, 2022), http://www.latimes.com/homeless-housing/story/2022-06-20/california-affordable-housing-cost-1-million-apartment [http://perma.cc/NKJ7-T6KP].
\end{itemize}
more than thirty percent of gross income on housing only to households earning in excess of $150,000. These housing costs are unaffordable to even the state’s wealthiest median income county (Santa Clara), and provides no wealth-creation pathway commensurate with homeownership for working families. Simply, the state’s housing policy preferences are not affordable, even for the state’s middle-class residents, like teachers, nurses, firefighters, and welders. State housing policies favor the wealthy, and taxpayer funded high-cost housing for the affordable lottery winners among the poor.

B. CEQA v. Agency Plans and Ordinances Allowing More Housing

The second most frequent target of anti-housing lawsuits is land use and zoning code approvals that make it easier to build more housing, including apartments, as required by state laws like SB 375, RHNA, and AFFH. Under SB 375, two of the state’s regional climate plans were targeted by CEQA lawsuits: these plans collectively allowed the two challenged regions (Bay Area and Sacramento) to accommodate a minimum of 594,688 new housing units as required by state law. Cities and counties are also required to update their local land use plans and zoning codes to accommodate RHNA-mandated housing allocations. The General Plan Housing Element approved by the City of Los Angeles, which was required to accommodate 456,643 housing units, was also targeted by a CEQA lawsuits. More than one million planned housing units were challenged in just eight of the 514 CEQA lawsuits filed during the study period. Just three CEQA lawsuits launched against agencies during our study period challenge more than one million new homes required to be built under state laws other than CEQA. Typical CEQA practice requires a housing project applicant to fund all compliance and defense costs, which increase housing costs for soon-to-be residents. Housing lawsuits filed against agency approvals of plans and ordinances that allow more housing are paid for by

---

89 For a primer on California’s regional housing needs assessment, affirmatively furthering fair housing civil rights law, and local housing element update laws, from YIMBY Law, see RHNA & Housing Elements, Explained, YIMBY ACTION, http://static1.squarespace.com/static/5fcea2bac5ab4f3059545081/t/603808983de3d40c40cb746a3f1614284956326/RHNA-Housing-Elements-Explainer.pdf [http://perma.cc/GT4C-WALE].
taxpayers, from the same pool of funding that pays for city parks, libraries, personnel, and other expenses.\(^90\)

We tabulated for this housing category only agency-approved plans and ordinances that allowed more housing units to be constructed. In Part III (CEQA v. Everything Else), we included challenges to agency approvals of plans, regulations, and ordinances that do not allow new housing—and, in some cases, actually make housing economically infeasible, even when the housing at issue fully complies with local General Plans, zoning, and regional SCS climate plans.

For example, one lawsuit sought to force San Diego County to use a methodology for demanding CEQA mitigation of “vehicles miles travelled” ("VMT"), the newest category of CEQA impact included in the CEQA Guidelines in the closing hours of the Brown Administration in 2017.\(^91\) VMT is measured only for miles driven by people in cars, mini-vans, and pickup trucks. Under CEQA, a housing project’s “VMT impact” is calculated by estimating how many miles these vehicles will be driven by construction workers, and then by future residents, guests, delivery and repair services, etc., over a thirty year home occupancy period. CEQA’s new VMT impact is separate from air pollution and greenhouse gas emission, which have long been estimated based on VMT. Even an all-electric car has the same VMT impact as a smog-belching 1970 Cadillac.\(^92\) Although

\(^90\) Cal. PUB. RES. CODE § 21089(a), (c) (Deering 2022) (stating public agencies may charge applicants to prepare CEQA documents); Att’y Gen. Bill Lockyer, Cal. Dep’t Just., Opinion Letter on Municipal Authority to Demand Indemnification from Third Party Lawsuits from CEQA Applicants (Feb. 4, 2002) (confirming municipalities may demand indemnification from third party CEQA lawsuits from applicants). But see Bryan Wenter, Game Changer: Public Agency Cannot Mandate Payment of Attorney Fees Under Indemnity Agreement Without Specific Statutory Authority, JDSUPRA (Mar. 9, 2021), http://www.jdsupra.com/legalnews/game-changer-public-agency-cannot-4713368/ [http://perma.cc/5CV5-Q9GB] (last visited Mar. 28, 2023); see also Stephen Kostka & Michael Zischke, PRACTICE UNDER THE CALIFORNIA ENVIRONMENTAL QUALITY ACT § 2.9 (2d ed. 2022) (noting that the Council on California Competitiveness, California’s Jobs and Future concluded that “[t]he higher cost of some EIRs often reflects the likelihood of litigation rather than the degree of environmental damage associated with a particular project”).


adding VMT to CEQA was originally justified as a GHG reduction mandate to address climate change, the correlation between VMT and GHG was substantially eroded following California’s mandated phase-out of internal combustion vehicles in favor of electric vehicle (“EV”) technology, along with its mandated transition to a 100% renewable energy grid. VMT’s regulatory promoters at the OPR pivoted to assertions that reducing vehicular use would improve water quality based on avoided use of vehicular brake pads, and reduce conventional air pollutants causing smog (although EPA had concluded that tailpipe emissions of smog had been reduced ninety-nine percent as of 2016). VMT promoters also extolled the health benefits—"wellness"—of people living in high density, walkable neighborhoods near existing job centers, and asserted that achieving these health outcomes through newly-defined impacts was within CEQA’s public health protection scope. In practice, CEQA’s VMT focus converted to imposing higher CEQA mitigation costs—to mitigate “VMT impacts”—on new housing built outside “transit priority areas” (“TPAs”). A TPA is defined as the half mile radius around high frequency bus stops, which in turn must have a minimum of four separate buses providing service for each weekday morning and afternoon peak commute as well as minimum evening and weekend service, or are near commuter rail stations or ferry terminals. Since these are the suburban scale and master planned community neighborhoods where most housing—especially for homes that middle income Californians can afford to buy—are located, and continue to be


95 CAL. PUB. RES. CODE § 21099 (West 2023).

96 Id.
built at attainable prices.\textsuperscript{97} TPAs were defined as neighborhoods located within a half mile of a metro or rail station, commuter ferry, or high frequency bus routes that met minimum standards of one bus every fifteen minutes during morning and evening peak commute hours (eight drivers and four buses for each routes), with further minimum service thresholds for weekends and evenings.\textsuperscript{98} Public transit generally, and TPAs more specifically, occur on a minute fraction of California’s one hundred million-acre footprint.\textsuperscript{99}

A report prepared for the Southern California Association of Governments (“SCAG”) region illustrates how rare TPAs are in California.\textsuperscript{100} The SCAG region includes all Southern California counties and cities except those in San Diego county—transit accounts for only 5% of total trips in the region.\textsuperscript{101} However, both transit service and actual transit utilization occur in just a fraction of the region: 82% of these transit trips occurred in Los Angeles County, 8% in Orange County, and the remaining 10% distributed between Riverside, San Bernadino, Inyo, and Ventura counties.\textsuperscript{102} Most of the region’s transit commuters live on only 1-3% of SCAG land, located overwhelmingly in Los Angeles county, and commuter transit use correlates to higher quality reliable transit service provided by TPAs.\textsuperscript{103} Los Angeles county is approximately 4,750 square miles—3% of the county is 1,410 square miles.\textsuperscript{104} The SCAG region is more than 38,000 square miles – transit ridership is either unavailable, or infrequent, for the vast majority of the region.\textsuperscript{105}

A spacial map of TPAs in the state’s next most populous region, the San Francisco Bay Area, further illustrates the mismatch between the amount and distribution of land in TPAs in this nine-county 6,900 square mile region:

\textsuperscript{97} See CEQA Transportation Impacts (SB 743), GOVERNOR'S OFF. OF PLAN. & RSCH., http://opr.ca.gov/ceqa/sb-743/ [http://perma.cc/2MCL-NFNB].
\textsuperscript{98} See id.
\textsuperscript{99} See Hernandez, Green Jim Crow, supra note 12.
\textsuperscript{101} Id.
\textsuperscript{102} Id.
\textsuperscript{103} Id.
\textsuperscript{104} Id.
\textsuperscript{105} Id.
Only San Francisco, a forty-nine-square mile peninsula, is almost entirely a TPA. The region’s other TPAs largely follow corridors on both sides of the cities fronting San Francisco Bay, and add mile-wide donuts in the downtowns of some larger cities and towns. Encouraging higher density housing in TPA—for those who can afford $3,500 monthly rents or condos over $1 million—may reduce VMT, but so does remote and hybrid work facilitated by high quality broadband as we learned during COVID. Using CEQA to add housing costs—VMT mitigation costs—in the vast

---

majority of the region that is located outside the TPA circles and lines just adds housing cost burdens to the region’s notoriously costly housing market, where high housing prices have driven more than 100,000 daily commuters outside the region into the adjacent San Joaquin county.\footnote{Commute, San Joaquin Council of Governments California, http://www.sj cog.org/230/Commute [http://perma.cc/Q6PB-9SEK] (last visited Apr. 15, 2023).} Only 1.5% of San Joaquin county residents take public transit to work or school.\footnote{Id.} Driving housing costs further higher in the the non-TPA Bay Area to “mitigate” VMT impacts makes it even less likely that a Bay Area worker can afford a home near work.

At a statewide scale, VMT is even more punitive: California is about 163,695 square miles.\footnote{See How Big is California, SpareFoot, http://www.sparefoot.com/moving/moving-to-los-angeles-ca/how-big-is-california/ [http://perma.cc/4BFL-PDMX].} As shown by the SCAG study, TPAs are scarce to non-existent in counties and cities outside already-urbanized, higher density locations.

Under the OPR VMT Guidance, which was new housing projects in cities are currently required to have VMT that is fifteen percent lower VMT than the “average” VMT for the city. For housing and other projects subject to CEQA review and discretionary approvals in the unincorporated areas of counties outside city boundaries, OPR has directed that VMT should be fifteen percent lower than the combined VMT average of all cities plus the unincorporated county.\footnote{See David Taub, State Law Could Push Middle Class Out of Housing, GV Wire (Mar. 5, 2020), http://gvwire.com/2020/03/05/state-law-could-push-middle-class-out-of-housing/ [http://perma.cc/3HSB-WSA8].} This methodology, which has been most heavily litigated in San Diego county, means that even higher density housing in the county cannot meet this standard because so much of the population lives in the cities along the coast where driving distances are shorter and, in some locations, TPAs do exist and provide meaningful transit services.

The fallacy of the metric, which was invented by a consulting firm that has subsequently earned many millions of dollars selling its VMT analytical services to cities and counties statewide, is that a new apartment or home built in an existing neighborhood presents occupants with the exact same suite of transportation needs and solutions as their new neighbors in the existing housing next door. The class- and race-based discrimination inherent in this new CEQA metric is that high wealth neighborhoods, with better schools and parks, have little or no public transit – and no TPAs. State housing laws, like Affirmatively Furthering Fair
Housing, mandate distribution of new housing throughout the community, including, for example, adding apartments and affordable housing in “high resource areas.”\textsuperscript{111} This is based on decades of civil rights studies demonstrating that poor and minority residents of high opportunity neighborhoods achieve higher educational attainment and income levels than those who grew up in poor neighborhoods with poor schools and fewer parks and other amenities.\textsuperscript{112}

San Diego County, which is required to accommodate 6,700 new housing units,\textsuperscript{113} was sued under CEQA to require strict adherence to non-regulatory state “guidance” on how VMT should be addressed under CEQA—one of the lawsuits included in this Study. The County Board of Supervisors, the majority of whom are aligned with open space and urban limit line advocates, opposed to new development in the County, directed staff to fully enforce the state VMT guidance. The result: for the three quarters of 2022, the County approved about 60 housing units per month. Once the VMT CEQA mitigation regime became effective in September 2022, permitting dropped to 8 units per month. In testimony provided on March 1, 2023, County staff reported that VMT mitigation fees, which can cost $50,000 or more per apartment, are likely making much of the housing outside of TPAs economically infeasible.\textsuperscript{114} The imposition of VMT as a CEQA impact effectively negates much of the County’s state-mandated and approved Housing Element, which is required to equitably distribute housing across the County (including within the County’s high opportunity but high VMT neighborhoods), as well as provide for housing solutions affordable to the region’s residents, including aspiring


\textsuperscript{112} See, e.g., How do Neighborhoods Affect Economic Opportunity?, EQUAL OPPORTUNITY PROJECT, http://www.equality-of-opportunity.org/neighborhoods/ [http://perma.cc/5Z4M-CFRM] (last visited Mar. 26, 2023) (“Studying the experiences of seven million children who moved across areas while growing up, we document that every year of exposure to a better environment improves a child’s chances of success.”).


\textsuperscript{114} Email from Matt Adams, BIA of San Diego to San Diego Board of Supervisors (Mar. 1, 2023) (on file with author).
homeowners seeking to close the racial and generational wealth gap created by California’s anti-homeownership/housing policies.

Bus ridership was crashing in Southern California (and the rest of the nation) even before COVID.\textsuperscript{115} Since COVID, and with the advent of remote and hybrid work patterns, bus ridership in much of the state has yet to recover to even sixty percent of its pre-COVID levels.\textsuperscript{116} As noted above, CEQA VMT mitigation costs such as $50,000 per apartment (and more for the cost of a home) even though future residents will use the same transportation options as their next-door neighbors is particularly punitive and disproportionate as a climate strategy, particularly since new homes must be built to stringent Green Building Code compliance standards and for example will use far less water and energy than the existing homes\textsuperscript{117} occupied by the legacy residents of these “nice” neighborhoods.\textsuperscript{118}

VMT is one of the environmental/climate redlining metrics discussed in \textit{Green Jim Crow}.\textsuperscript{119} It is also an example of an anti-housing CEQA metric embraced by environmental agency staff and anti-housing NIMBYs and advocates to continue to structurally embed in CEQA anti-housing mandates that undermine housing and civil rights laws, like Affirmatively Furthering Fair Housing.

C. CEQA v. Homeownership for Middle Income Families

In third place are neighborhood community housing projects, which generally include a mix of single family homes as well as townhomes or condominiums, “accessory dwelling units” (“ADUs”), either in the form of backyard cottages or granny flats located within the main home structure, and small project subdivisions of fifty or fewer homes. These neighborhood-scale community housing projects also include, or are proximate to, parks and retailers, and may include new elementary schools, fire stations, or other public services. The largest of this home type is a “Master Planned Community” (“MPC”), which is planned at a larger scale and typically includes several thousand housing units in different

\textsuperscript{116} See Epstein, supra note 113.
\textsuperscript{117} CAL. CODE REGS. tit. 24, part 11 (2022).
\textsuperscript{118} See Adams, supra note 114.
housing types at different levels of affordability, as well as new infrastructure, and—for purposes of our studies—is large enough to include a new high school. These are all projects that typically primarily include for-sale homes of varying sizes — given the importance of homeownership as part of the California Dream of working families — and help produce sufficient new homes to close the racial wealth gap created by more than a century of racial redlining that persisted into the 2008 recession with predatory loans and foreclosures that disproportionately targeted homeowners of color. As compiled by affordable housing producer Habitat for Humanity, homeownership has long been recognized as the nation’s most successful pathway to build inter-generational wealth, as well as housing stability and other civic benefits such as higher educational attainment, higher rates of community volunteer activities and voter participation, etc. The wealth gap between renters and homeowners is staggering: a September 2020 report from the Federal Reserve found that, on average, a homeowner had forty times more wealth than a renter. A legacy of racial discrimination, which persisted into and beyond the 2008 Great Recession’s foreclosure crisis, has resulted in far lower homeownership rates for California’s Black and Latino families — and in 2022, fewer than one in five Black or Latino families in California could afford to own a median priced home. California environmental policies favor high density urban rental apartments that are unaffordable, and strongly disfavor building new homes on lower cost land, in lower cost locations, with lower cost structures that are actually affordable for either purchase or rent by working families. California’s climate-based policies double down on these NIMBY environmental policies, expressly acknowledging the disproportionately higher economic burdens

---

120 This infrastructure may include: public services, like new fire stations and schools; job-creating commercial, retail, and institutional uses; renewable energy; and other sustainability features.


placed on median and lower wage (more likely Latino and Black) households while favoring wealthier (Whiter and Asian) households. Neither Californians nor state elected leaders have voted to end attainable homeownership or kill the California Dream for anyone but the wealthy, but state leaders and bureaucrats have enthusiastically embraced or enabled policies that have caused exactly this outcome for decades, including most recently by voting to approve the CARB climate plan’s $5.3 billion wealth transfer scheme to increase climate cost for households making $100,000 or less while reducing $5.3 million in climate costs from higher income households. California has the second worst homeownership rate in the nation, morphing the California Dream for a fading Baby Boomer legacy generation of median income households to lifelong renters for all but their wealthiest successors.

The underlying policy debate is an epithet: “Sprawl.” Californians despise “sprawl” as causing traffic gridlock, but disagree as to what sprawl actually is and where new housing should actually be located. On one end of the spectrum, because state housing and environmental laws mandate that new development be “green,” require dispersal of housing throughout communities under Affirmatively Further Fair Housing, and have not rescinded the civil rights and equity laws and regulations to make homeownership attainable to communities of color and middle class families of all colors, local governments and housing applicants continue to plan for and approve this housing and families continue to save for, and buy, their first new home. On the other end of the spectrum, environmentalists—long committed to blocking development even on proximate urban lands and

---

127 See e.g., Lazo, supra note 124.
phasing out automobile use (even electric automobiles)—oppose single family and other lower density housing even when long-planned in existing cities needing workforce housing. At the most extreme end of this environmentalist spectrum are Malthusians who believe that California (and Earth) are at risk of reaching their holding capacity, and “de-growth” of California is necessary (fewer people overall, and no more housing growth except expensive, high density, small rental apartments in transit-dependent neighborhoods) to ward off climate change and the mass extinction of species.

In fact “sprawl” is generally used to refer to single family homes built in suburbs to satisfy consumer demand for “homes with more square footage and yard space” and avoid the “traffic, noise, crime, and other problems” of cities.” “Smart growth” emerged in opposition to “sprawl,” and promotes building new homes only by substantially increasing densities in existing cities and towns.

Smart-growth-only advocates underestimated voter resistance to density, the much higher cost (and reduced homeownership opportunities) of an all-densification urban limit line regulatory regime, and the continued desire by people to have more living and outdoor space away from the noise and bustle of high density cities for at least some portion of their life (e.g., when raising children).

In my view, neither sprawl nor smart growth have worked well: Baby Boomer battle lines that are decades old have resulted in massive housing shortages, obscene housing prices in the most “progressive” Green anti-housing political enclaves like San

---

131 See id.
134 David B. Reskinik, Urban Sprawl, Smart Growth, and Deliberate Democracy, 100 AM. J. PUB. HEALTH 1852, 1853 (2010).
Francisco and Marin county, and fragile or dysfunctional transportation, water, and energy infrastructure notwithstanding California’s exceptionally high tax and fee burdens. Like many “Zero Sum” debates promoted by partisan special interests, neither “sprawl” nor “smart growth” can provide solutions for the fact that California’s population is about twice as large today as it was when CEQA was enacted in 1970.136

Harvard University’s Education Department published an influential study of solutions for affordable and sustainable housing in Mexico (“Harvard Study”), which first suggests strategies for increasing infill density but then goes on to explain that “even in metropolitan areas with successful records of infill development, infill as a percentage of total area growth remains a minor portion of total growth” and “[g]reenfield development, or development on previously undeveloped sites, must be an equally important aspect of city-building in the 21st century if urban areas are to properly and adequately house new generations of city-dwellers.”137 As summarized on the next table, with information from the Harvard Study, Sustainable Greenfield Development—often referred to in practice as Master Planned Communities—substantially differs from “sprawl.”

<table>
<thead>
<tr>
<th>Characteristics of Urban Sprawl</th>
<th>Characteristics of Sustainable Greenfield Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low residual density</td>
<td>Higher overall residential density with a variety of housing types, not just single-family houses</td>
</tr>
<tr>
<td>Unlimited outward extension of new development</td>
<td>Outward extension of development is limited by numerous factors, including municipalities’ ability to provide infrastructure and services, open space preservation, and environmental protection considerations, etc.</td>
</tr>
</tbody>
</table>

138 Id. at 95.
| Spatial segregation of different types of land uses through regulations | Land use types are mixed and integrated, with town centers, office parks, and other employment and commercial centers easily accessible from residential areas |
| Leapfrog development (or development that leaps out onto new land, not connected to existing urban areas) | Contiguous urban expansion |
| No centralized ownership of land or planning of land development | Land development happens in accordance to well-defined plans or in cooperation among landowners |
| All transportation dominated by privately owned motor vehicles | Infrastructure and development supportive of many modes of transportation are created, including bus, rapid transit, bicycles, and pedestrians |
| Fragmentation of governance authority of land uses among many local governments | Governance of land use is coordinated among all municipalities in a region |
| Great variation in fiscal capacity of local governments | Commercial development is concentrated in nodes or town centers, serviced by a multi-modal transport network, not just roads for automobiles |
| Widespread commercial strip development along major roadways | Affordable housing is provided through a combination of an increased supply of housing, a variety of housing types, government requirements (like inclusionary zoning) and government programs, among others |
| Major reliance on filtering process to provide housing for low-income households. Filtering occurs when wealthier people move into new homes and low-income people move into the older and lower-quality houses left behind. |

My selection of this Harvard Study is intentional: the country lacks the wealth of California, and the study is designed to promote an equitable, as well as environmentally and financially sustainable, solution to an even more severe housing and poverty crisis. Mexico is getting wealthier, with job and income growth,

\[\text{139 The World Bank in Mexico, The World Bank GRP.,}\]
and advocates are seeking to use that wealth to promote a positive outcome for people—and the environment.

In citing the Harvard Study, I hope to, at least in part, bypass the fractious and pessimistic stand-off between strident anti-single family home environmentalists, and equally strident anti-densification environmentalists, who have used tools like CEQA to elevate legal procedure and process over solutions to our housing, infrastructure, and climate challenges. This stand-off, and the labor movement’s willingness to tolerate this stand-off, even as it hurts middle income labor union households the most, have mostly “preserved” the increasingly imperfect status quo (unless you are already a wealthy donor who owns a home).

The housing crisis would be much easier to solve (and the state would reduce its greenhouse gas emissions and save the planet) if only we had far fewer Californians. That is not a racially-just outcome—it just honors Boomer nostalgia for free-flowing roadways and climate catastrophists convinced that getting America’s lowest per-capita greenhouse gas emission state to “net zero” requires making the state unaffordable to all but its wealthiest residents (and their NGO and academic grantees).

In the world of CEQA lawsuits, Californians are losing and the Malthusians are winning: favored housing is unaffordable and sued under CEQA, disfavored housing is affordable and sued under CEQA, California’s population is decreasing, and CEQA lawsuits to block even planned and approved housing that meet all of California’s stringent green standards are the favored tool to achieve the anti-housing policy objectives of both Malthusians and environmentalists.

D. CEQA v. Students

CEQA lawsuits to block student dormitories were a major new target in this Study, even as colleges and universities have recognized that the absence of proximate, affordable student housing is causing massive harms such as homelessness, anxiety, and high drop-out rates.140 Housing insecurity also causes greater

---

140 See U.S. DEPT. HOUS. & URB. DEV., INSIGHTS INTO HOUSING AND COMMUNITY DEVELOPMENT POLICY 1 (2015); Michael Burke et al., How California is Responding to Dire Student Housing Shortage, EdSource (Sept. 28, 2022), http://edsource.org/2022/how-california-is-responding-to-dire-student-housing-shortage/678616 [http://perma.cc/KFT9-9Q4J]; see also Brief for The Two Hundred for Homeownership as Amici Curiae, Make UC a Good Neighbor v. The Regents of the University of California et al., Case No. A165451
harm to students of color, and students who are the first in their family to attend a four year college, just as they begin an educational journey that has in all past generations promised upward mobility and higher incomes.141

Rents that students pay to live in dormitories serve as a viable financing source to pay for the construction of new housing, allowing dorms to be built without triggering the need for tuition increases or budget cuts to other college programs. Because these dormitory projects are also generally required to pay higher wages to construction workers, similar to other public agency infrastructure projects, organized labor has not filed CEQA lawsuits against student dorms. A fierce and unapologetic constituency of literal NIMBYs campus neighbors has turned to CEQA: tens of thousands of new student beds are challenged in seventeen anti-dorm CEQA lawsuits filed during the Study Period.

Although reporting on the outcome of these CEQA lawsuits is beyond the scope of these studies (typically because the final resolution of CEQA lawsuits is not known for three to five years), it is noteworthy that UC Berkeley was the target of more of these anti-student housing lawsuits than any other campus. In one of several different CEQA lawsuit decisions, UC Berkeley was ordered to admit three thousand fewer undergraduates, a trial court decision that the California Supreme Court declined to review just a few days before student admission letters were scheduled to be mailed.142 The Legislature instantly stepped in, decrying the concept that students were “pollution” or “anti-environment”—but the enacted “fix” Legislation was exceptionally narrow,143 and did nothing to block pending anti-university CEQA lawsuits. For the first time in CEQA’s fifty-three-year old history, an appellate court had determined that the “social noise” of future student occupants of future student dormitories was indeed an “environmental impact” requiring evaluation and “all feasible mitigation” under

---


CEQA.144 This decision creates a broad path for future lawsuits against housing for teenagers (music!), families (babies who cry!), and others who do not live the quiet life of the retirees who chose to purchase a home next to the University of California’s oldest campus and now want it to be “QUIETER! Gosh darnit!”

E. CEQA v. Old People

More CEQA lawsuits were filed against housing for the elderly than housing for the homeless. On even the most benign scale of housing “impacts” to the environment, senior housing ranks at the rock bottom: it generates far fewer traffic trips overall and during commute hours, it generates no “students” to crowd parks and schools, and “social noise” impacts of future residents are likely limited to the volume setting of an individual TV. Building senior housing also plays an outsized role in helping alleviate the housing crisis: seniors are most likely to move from existing single family homes, making those homes available for purchase by younger families who are otherwise renting, which in turn creates a new unit on the rental market.145

F. CEQA v. Homelessness

The Legislature enacted numerous CEQA exemptions designed to streamline the construction of shelters and other housing for those experiencing homelessness, including a statutory exemption from CEQA for converting hotels and motels into housing for unsheltered residents. As reported by scholars at UC Berkeley, this worked:146 Project Roomkey provided temporary housing to 22,000 people as of the end of 2020,147 and Project Homekey has funded 12,676 hotel conversion permanent housing units.148 To the legions

144 See Make UC a Good Neighbor v. Regents of Univ.of Cal., 384 Cal. Rptr. 3d 834, 850, 861 (Ct. App. 2023).
of anti-housing CEQA defenders, however, a statutory exemption just sets the legal framework for CEQA lawsuits asserting that the exemption does not (or should not) apply.

G. CEQA v. Single Family Homes/Casitas

The final noteworthy category of anti-housing CEQA lawsuits involves single family home projects, including new homes on existing lots and home remodels. The most notorious of these lawsuits languished in court for eleven years, including two trips to the California Supreme Court, in a gadfly v. homeowner dispute over the rebuild of a single family home on a single family lot in Berkeley.\(^\text{149}\) The home rebuild was unanimously supported by the Berkeley Planning Commission and City Council.\(^\text{150}\) There is a longstanding regulatory exemption (called “categorical exemption” in CEQA-ese”) finding that building a single family home on a single family lot does not cause environmental impacts warranting further study under CEQA.\(^\text{151}\) A “community activist” sued anyway, decrying the size of the home and asserting that the Berkeley Hills were susceptible to landslide risks (they are, and buildings must meet stringent standards to protect against landslide risks).\(^\text{152}\)

The same CEQA housing opponent lawyer in Berkeley Hillsides sued on behalf of NIMBY neighbors to block another single family home rebuild in the small Marin County community of San Anselmo\(^\text{153}\) (median home price, $2.1 million).\(^\text{154}\) The neighbors unsuccessfully argued that the home and neighborhood were entitled to historic preservation status, in a year-long dispute


\(^{151}\) CAL. CODE REGS. tit. 14 § 15303(a) (2023).


that included expert reports and contested hearings, before the elected city council approved the project. The neighbors then sued under CEQA; many months later, their lawsuit was found to have no merit in an exceptionally detailed trial court decision. The neighbors then filed two appeals, offered to drop their then-pending appeal only if the homeowner agreed not to seek to recover the modest court costs the neighbors would have otherwise had to pay, then waited until the last day to drop their appeal even when their cost-avoidance request was rejected. Their courtroom tactics cost another year’s delay during COVID.

In an unusual twist to the normal CEQA lawsuit story, where the losing NIMBY-side’s lawyer—having cost the project applicant years and hundreds of thousands of dollars—simply slips away quietly and with no financial consequences to the next anti-housing CEQA lawsuit, the homeowner applicant sued the CEQA lawyer for engaging in “malicious prosecution” in bringing a meritless lawsuit alleging that the city had violated CEQA and land use law and then manipulating the appellate process to avoid court costs. The target of the malicious prosecution lawsuit has herself argued multiple cases before the California Supreme Court and has been hired by the state judiciary to teach CEQA to state judges in its mandatory CEQA education program. The appellate court reviewing the malicious prosecution issue, in the context of the lawyer’s motion that the lawsuit should be dismissed as an “Anti-SLAPP” (strategic lawsuit against public participation) infringement of her protected Constitutional right to engage in the challenged conduct, found that the lawyer’s conduct was indeed grave enough that it demonstrated “a probability of prevailing” on the malicious prosecution claim, meeting all three required criteria:

155 See Jenkins, 302 Cal. Rptr. 3d at 889–92.
156 See id. at 892, 894.
157 See id. at 895.
158 See id. at 892, 895.
159 Id. at 895–96.
The claims were without merit, as had been exhaustively explained by the trial court.162

There was no “probable cause” that the claims would prevail. On the land use claim, the appellate court was persuaded that a deliberate and highly misleading argument, and related record reference, about whether a standard was mandatory (petitioner wins) or permissive (discretionary, and petitioner loses) showed the absence of probable cause as to the existence of a meritorious claim. On the CEQA claim, the appellate court found that it was barred for failure to exhaust administrative remedies: a known, and jurisdictional, bar to filing a CEQA lawsuit. Further, the appellate court found the CEQA argument to be invalid even had the argument been timely made to the city because it was directly at odds with the Supreme Court’s decision in Berkeley Hillsides, the eleven-year CEQA anti-single family home rebuild saga described above that the same lawyer had litigated on behalf of a different anti-housing NIMBY, and lost.163

Most remarkably, the appellate court found that the lawyer had acted with “malice” based on the “subjective intent or purpose.” The Court noted:

Defendants’ failure to present the record fairly supports a finding they knew their claims were untenable, Defendants made misleading arguments, Defendants filed and swiftly dismissed the Writ of Supersedeas, and Defendants maintained their appeal for three months and offered to dismiss the appeal only if Plaintiffs agreed to waive any claim to fees and costs.164

While it may be tempting to dismiss this San Anselmo lawsuit against a single-family home rebuild with a tiny granny cottage in the backyard as an anomaly in CEQA lawsuits, it is, in fact, the entirely “unremarkable” pattern of CEQA lawsuits. Two of the most ardent defenders of the CEQA status quo, UCLA Law Professor Sean Hecht (who has since moved on to work for Earthjustice) and former Chief of Staff for the California Attorney General (under Jerry Brown) and current UC Davis Law Professor Rick Frank filed an amicus brief in support of the

---

162 See Jenkins, 302 Cal. Rptr. 3d 893.
163 See generally id.
164 Id. at 905.
lawyer sued for malicious prosecution in this case. Their amicus argued that this lawsuit was entirely “unremarkable,” that the outcome of CEQA lawsuits was massively uncertain, and the record showing what the court determined to be erroneous and misleading factual and legal arguments raised by the attorney in her court pleadings was simply a “ubiquitous” feature of CEQA lawsuits.

The appellate court reviewed this and other amici in its decision. The court specifically rejected arguments raised by amici that the lawyer was simply working to protect “the environment,” noting that “the Jenkinses’s situation has nothing to do with environmental protection and everything to do with the privacy and aesthetic design concerns of several of the Jenkinses’s neighbors.” The court also rejected amici arguments asserting that CEQA was critical to protecting disadvantaged communities of color, arguing that “the Jenkinses’s lawsuit has nothing to do with ‘disadvantaged communities,’ ‘underserved communities,’ ‘marginalized communities,’ ‘pollution,’ ‘human health consequences,’ or ‘urban decay,’ to name just a few of the topics raised” by amici. Instead, the court found “apt” the Jenkinses’ brief, which argued that this CEQA lawsuit:

involved a group of well-off, ‘NIMBY’ neighbors living in one of the most expensive zip codes in the country trying to prevent their fellow neighbor from rebuilding a decrepit and dangerous residence on their property because the neighbors were concerned about privacy the design aesthetics of the new build. It had nothing to do with significant or negative environmental effects under CEQA.

III. CEQA V. EVERYTHING ELSE (NON-HOUSING)

Although housing is the top target of CEQA lawsuits, at 39% of all lawsuits filed during the study period, 30% of CEQA lawsuits target public infrastructure and other non-residential community construction projects, 26% target non-residential private sector construction projects, and 5% target agency plans and regulations that do not approve specific construction projects, as shown in Figure 3.

---

166 Id. at 9, 11, 20.
167 Jenkins, 302 Cal. Rptr. 3d at 907.
168 Id.
We have categorized these into Non-Residential Public and Community Projects (Figure 4), Private Sector Non-Residential Projects (Figure 5), and Public Agency Non-Residential Plans and Regulations (Figure 6). Collectively, these non-housing CEQA lawsuits demonstrate the power of one CEQA lawsuit to thwart laws and decisions that would change the status quo, either by those seeking to preserve the status quo (NIMBYs or other incumbent stakeholders opposed to change), or by those seeking to leverage CEQA lawsuits for economic benefits (competitor or wage lawsuits). We discuss each in turn below.
A. CEQA v. Public Infrastructure, Public Services, Utility, and Renewable Energy Projects

These projects all involve agency decisions to authorize some physical change to the environment at a particular location.

1. CEQA v. Water Equity

By far the largest target of CEQA lawsuits in this public project category are agency decisions to manage or increase water supplies, as shown in Figure 4. Although climate change is routinely blamed for weather events, including droughts, California’s over 130 year
record of annual precipitation and droughts shows massive variability year-over-year, as published by the National Oceanic and Atmospheric Administration as part of their National Integrated Drought Information System and National Center for Environmental Information, and reprinted below:\(^{169}\)

\[\text{CALIFORNIA PRECIPITATION}\]

This historic pattern, which includes many decades not attributed to the post-1960 decades most associated with higher carbon content in the atmosphere and climate change, demonstrates that California cannot rely on any “natural” condition to provide itself with an adequate “natural” year-round water supply in the right locations based on rainfall. California’s major population centers have always relied on imported water to meet demand, with its oldest and wealthiest cities in the Bay Area, importing most of its water from a dam built in a canyon of Yosemite National Park\(^{170}\) and its wealthiest cities in Southern California, importing water from the Eastern Sierras, the


Sacramento Bay Delta, and the Colorado river. Water management infrastructure facilities—storage, conveyance, treatment, and distribution—are absolutely critical public infrastructure all along the California coast, with, for example, a new generation of desalination plants installed or approved (and delayed) with CEQA lawsuits, along with recycled, reclaimed, and other local water storage, levee, and environmental enhancement, water quality, and other water management projects approved—and sued under CEQA in cases included in this dataset.

Simply, California needs water storage and conveyance solutions for both normal and drought years. There is nothing new about the need to manage water: from localized irrigation built by early human farmers to the famous aqueducts of the Roman Empire and beyond, access to reliable water supplies have been a core societal need. In California, public agencies are charged with meeting that need—either directly through water agencies, or indirectly through the regulation of private water utilities and companies. Massive federal and state water projects, and smaller-scale local water projects, were partially or mostly completed—sometimes nefariously—for more than a century after statehood, from San Francisco damming a portion of Yosemite National Park to Los Angeles draining much of the Owens Valley east of Yosemite.


173 See CAITRIN CHAPPELLE, ELLEN HANAK & ANNABELLE ROSSER, PUBLIC POLICY INSTITUTE OF CALIFORNIA WATER POLICY CENTER, PAYING FOR CALIFORNIA’S WATER SYSTEM (2021).


By the early 1970s, the state’s first twenty million overwhelmingly white residents built and prospered from a world-class network of reservoirs and aqueducts indispensable for living in a desert with massively variable precipitation. Since then, the state added nearly another twenty million people, almost entirely Latino, Asian and other minorities. Starting in the 1970’s, though, politically influential white activists have blocked virtually all major new water storage and distribution system improvements, including those that are essential for providing new, less affluent families with affordable, reliable water.

The state’s two biggest water projects (the federal and state projects) were never completed, and remain stressed by competing water demands, insufficient water supplies, and increasingly fragile physical facilities like deteriorating levees in the Delta and deteriorating dams—both at risk of catastrophic failures in heavy storms or

---


179 See Johnson et al., supra note 176.


182 See State Water Project, WATER EDUC. FOUND., http://www.watereducation.org/aquapedia/state-water-project (last visited Mar. 2, 2023) (“The SWP originally was conceived as a much larger project, but only its first phase was completed.”); see also Hernandez, In the Name of the Environment II: 2013-2015, supra note 1.

earthquakes\textsuperscript{184} which would result in multi-month water supply and delivery shortfalls.\textsuperscript{185} Voter-approved funding to increase water availability and reliability (including water storage facilities\textsuperscript{186}) has instead been spent primarily on open space acquisition and conservation and agency/consultant staff, with no appreciable increase in reliable and affordable water supply deliveries.\textsuperscript{187} Regulatory hurdles to transferring water from willing sellers to thirsty buyers remain daunting,\textsuperscript{188} and water conveyance facilities are also stressed in many parts of the state.\textsuperscript{189} Non-partisan reports by public agencies and academics have estimated that over one million Californians, primarily poor and non-White, do not have access to safe drinking water from the taps in their homes.\textsuperscript{190} Although the Legislature enacted a “Human Right to Water” law in 2012,\textsuperscript{191} none of the water storage facilities approved by the voters are anywhere near approval, and have instead been sidelined by ever-


\textsuperscript{188} See, e.g., \textit{Water Transfers Program}, CAL. WATER BDS. (Mar. 6, 2023), http://www.waterboards.ca.gov/waterrights/water_issues/programs/petitions/transfers.html [http://perma.ca/ZRW3-4KHN]; SOAR WATER TRANSFERS ACTION TEAM, PROPOSALS TO STREAMLINE WATER TRANSFERS (2014). Both materials reveal that the transfer process is lengthy, and attempts at streamlining this process are ongoing.

\textsuperscript{189} See \textit{ELLEN HANAK ET AL., MANAGING CALIFORNIA’S WATER} 151–52 ( 2011).


\textsuperscript{191} See CAL. WATER CODE § 106.3 (West 2013).
escalating costs and ongoing opposition. Although the water system failures in Flynt, Michigan made headlines—and resulted in criminal charges against those responsible—in 2022 the California Legislative Analyst’s Office concluded that the key state water agency in charge of these issues was not acting with any urgency to solve state water equity, reliability, safety and affordability legal mandates. In short, beginning with the era of modern environmental laws, including CEQA in the 1970’s, California stopped building the water infrastructure needed for its growing population and highly and persistently erratic rainfall patterns.

Today, even moderately dry conditions (sixty to eighty percent) of “normal” rainfall years are enough to trigger yet another emergency declaration, demand for water use cutbacks, and panicked polls ranking water supply as the top environmental concern instead of wildfires and climate change. CEQA requires that significant new housing projects demonstrate sufficient water supplies during normal, dry, and multiple dry-year periods. These housing projects cannot be built without adequate water supplies. Blocking “new” water supplies is a potent anti-housing tool that has long been used in infamously NIMBY green communities like Marin County to block new housing.

---

193 See PETEK, supra note 190, at 11–12.
197 See Dan Walters, Marin County’s Guerilla War Against Housing, CAL MATTERS (May 31, 2021), http://calmatters.org/commentary/2021/05/marin-county-housing-water-quota/ [http://perma.cc/YSR7-CJDW].
As shown in Figure 4A, below, CEQA lawsuits are most frequently used to target agency decisions to allocate existing water supplies among users. Such decisions typically increase costs to existing users and divert a greater percentage of water deliveries to the environment to support fish and other habitat values. In the lawsuits reviewed for this article, incumbent water consumers, including cities such as San Francisco and farmers in the Central Valley, sued to block new agency decisions to allocate (or change allocations) water. Figure 4A shows that the next most likely water project to be targeted by CEQA lawsuits are those that would increase the availability of water supplies to people. In the lawsuits reviewed for this article, these included projects to treat and use recycled water, desalination projects, groundwater basin recharge projects, and projects to convey water to people by pipeline. Also, famously anti-growth advocates in Monterey County have “enjoyed” a full moratorium on the construction of new housing (even granny flats) based on an insufficient water supply. 198 It appears from our study that every significant water supply augmentation project in that region is sued under CEQA. A small category of flood control projects is sued, including floodwater management that would increase groundwater storage or stormwater use, along water quality improvement projects and an emerging new climate category of sea level rise projects. The sea level rise climate debate pits those advocating for a “managed retreat”—abandonment of shoreline infrastructure and development—to those advocating for engineered solutions like sea walls to protect against storms and sea level rise. 199


Parks, schools, and streets make up the other more significant categories of these public and community projects, comprising just under ten percent of total CEQA lawsuits filed in this category.

2. CEQA v. Streets and Sidewalks

Figure 4 shows that streets and sidewalks are sued more under CEQA, typically for projects that remove street parking or trees to make way for bike paths, bus lanes, or “complete” streets that slow down traffic and promote pedestrian use. The “complete street” program was intensely criticized for narrowing the four-lane highway through the town of Gold Rush mountain community of Paradise to a two-lane road to facilitate downtown “walkability”—a project that left the town with woefully insufficient evacuation capacity, which in turn contributed to the catastrophic death toll for the Paradise wildfire.200 Bike paths

200 See Paige St. John et al., Paradise Narrowed Its Main Road by Two Lanes Despite Warnings of Gridlock During a Major Wildfire, L.A. TIMES (Nov. 20, 2018), http://www.latimes.com/local/california/la-me-ln-paradise-evacuation-road-20181120-story.html [http://perma.cc/UJ3Y-YLKR]. The newest anti-housing tool, currently most often used to block funding access and approvals of affordable and workforce housing in rural and resort areas, is wildfire risk of the scale that engulfed Paradise and other forested communities. Expert foresters have repeatedly cautioned that more than a century of forest mismanagement, which ended the sustainable forest conditions maintained by burns every ten years or so, coupled with predictable drought conditions,
also continue to be contentious targets of CEQA lawsuits, notwithstanding various partial legislative CEQA exemptions. A bike path resulting in the closure of a full lane of bridge highway traffic to accommodate a handful of Marin County daily bike riders, while delaying tens of thousands of daily workforce commuters into Marin County (where they cannot afford to live), lengthened commuter time for workers like teachers, and added to vehicular emissions stacked in disadvantaged communities outside of Marin County; the four-year pilot run of this bike path was not challenged in a CEQA lawsuit, but is likely to be targeted if made permanent.

3. CEQA v. Schools

Figure 4 demonstrates that projects regarding K-12 schools and colleges show the same pattern of NIMBY/incumbent status quo defense, and economic use, of CEQA. About two-thirds of school projects challenged student dorms (and are included in the anti-housing CEQA lawsuit challenges included in Figure 2, and thus excluded from Figure 3). The next biggest CEQA lawsuit targets are charter and religious schools (often opposed by public school parents and teachers), improvements to public school playfields (opposed by NIMBYs), and an only-in-San-Francisco COVID story of a CEQA lawsuit to block obliteration of historic murals painted by a socialist artist during the Great Depression in a San Francisco High School following a controversial vote by now-recalled School Board Members who asserted the mural


202 See CAL. PUB. RES. CODE § 21080.20 (2020); CAL. PUB. RES. CODE § 21080.25 (2020).

“glorifies slavery, genocide, colonization, manifest destiny, white supremacy [and] oppression.”

4. CEQA v. Parks/Trails

CEQA lawsuits are the favored tool used by passionate advocates for change, or not, to California’s parks and open space. Most park litigators are seeking to limit public use, or access, to parks (and block trails). The most productive source of CEQA lawsuits in all of California history is a narrow slice of Los Angeles land ending on the ocean immediately south of Marina Del Rey, where Howard Hughes built his aerospace empire, including the Spruce Goose. More than thirty CEQA lawsuits were filed over more than two decades to block development on the now-completed Playa Vista residential and office project west of Lincoln Boulevard, and the coastal strip from Lincoln to the ocean was required to set aside permanent open space including coastal wetland restoration in the Ballona Wetlands. CEQA lawsuits against this property have been a staple in all of our CEQA lawsuits, and this is no exception, as multiple advocacy groups filed CEQA lawsuits against the California Department of Fish and Wildlife over the management of this Ballona Wetlands open space.

B. CEQA v. Non-Residential Private Projects

The agency approvals in this category are for private sector projects that involve a physical change to the environment in a particular project location, as shown in Figure 5.


207 See, e.g., Hernandez, In the Name of the Environment I: 2010-2012, supra note 22, at 21 n. 50, 41 n. 103, 46 n. 161; Hernandez, In the Name of the Environment II: 2013-2015, supra note 1, at 69 n. 147.
1. CEQA v. Cannabis

In 2016, fifty-seven percent of California voters decided to legalize cannabis for adult, non-medical use. Proposition 64 established a comprehensive and ambitious program to tax and regulate the cultivation and sale of cannabis. However, it did not make growing or selling cannabis permissible uses in

---


California’s hundreds of cities and fifty-eight counties, nor did it regulate (for air pollution or water quality purposes) cannabis facilities. In the most remarkable new pattern to emerge from our earlier two CEQA studies, cannabis-related projects exploded into the second most likely to be targeted non-residential project in a CEQA lawsuit. Cannabis growers, retail outlets, and agency regulations and ordinances applicable to cannabis were equally likely to be targeted in CEQA lawsuits. Illegal cannabis operations, which do not obtain agency authorizations, are not sued under CEQA. State cannabis tax revenues are a fraction of what was promised by legalization advocates, the price of cannabis has plummeted since adult personal possession became fully legal, and sales are far less likely to be targeted by law enforcement. The cannabis regulatory framework was recently revised in an attempt to achieve more of the revenue and other objectives promised to voters in the legalization initiative.

2. CEQA v. Warehouse/E-Commerce

Figure 5 shows that warehouse projects are the most likely target of non-residential private sector projects to be sued under CEQA. Anti-warehouse CEQA lawsuits have evolved over a multi-year trajectory that displays the broad range of CEQA litigation status quo defenders.

Warehouse projects are relatively easy to assemble, with most of the work performed by laborers. In the first round of warehouse CEQA lawsuits, a union representing laborers would sue and settle with a “Project Labor Agreement” (“PLA”) in which the warehouse applicant would agree to use union members, and pay union wages and benefits, for warehouse construction. Another labor CEQA litigant has been a union representing truck drivers, prompted in part by a national

---

210 See id.
212 See id.
Amazon campaign\textsuperscript{215} and a concurrent explosion of e-commerce supercharged during COVID.\textsuperscript{216}

Another round of CEQA warehouse litigants, which continues to grow, has been a coalition of advocacy groups and local officials arguing that warehouse projects near residences (or that use roads near residences) are causing localized adverse air pollution conditions causing disparate harms to disadvantaged communities. The state’s Attorney General has aligned with groups focused on localized air pollution and other impacts.\textsuperscript{217} One less reported cause of the massive increase in warehouse facilities in Southern California is the fact that the ports of Long Beach/Los Angeles (the nation’s largest by cargo volume) is underserved by rail transport at the ports, and highly reliant on trucking.\textsuperscript{218} A multi-modal cargo facility that would have expanded rail capacity was blocked, in part by CEQA lawsuits included in our earlier studies, and a modified facility has been newly proposed but remains in the EIR process (pre-litigation).\textsuperscript{219}

California has the largest number of truck driver jobs by state.\textsuperscript{220} The Inland Empire (east of Los Angeles and Orange Counties) the state’s fastest growing economy, is itself more populous than half of U.S. states,\textsuperscript{221} and presently has a high percentage (more than 78%) of residents who do not have bachelors’ degrees.\textsuperscript{222} A recent Brookings Institute Report found

\textsuperscript{215} See id.
that a high percentage (about 40%) of Inland Empire residents were challenged in making ends meet each month, and that the logistics industry (including warehousing and trucking) was the region’s fourth largest employer (102,553 jobs). 223

A recent high profile clash pitted air quality regulators advocating for a transition to presently available ultra-low polluting fossil-fuel truck fleets to achieve air quality standards against climate and environmentalist advocates demanding a transition to an all-electric truck fleet, 224 which for heavy duty trucks is a decade or more away from being commercially available. The air quality regulators lost—currently unavailable EV heavy duty trucks are mandated by CARB 225 and older trucks will remain in service until new technology becomes commercially available.

Legislators, meanwhile, are seeking to ban warehouses and/or truck routes to and from warehouses in much of the region. 226 Into this policy, economic, equity, and environmental scrum marches the CEQA lawyers, including an infamous group which pursues a “sue and settle” CEQA business plan and was unable to identify, in a sworn deposition, that they have spent any of their CEQA settlement dollars on any identified environmental improvement projects. 227

3. CEQA v. Renewable Energy

Perhaps nothing better highlights CEQA’s antiquated, anti-environmental rigidity than its use against renewable energy projects, which the state’s climate laws and policy demand be built at an unprecedented scale and pace to avoid planetary catastrophe. 228 All but one of the energy projects sued during the Study Period was for renewable energy not generated from fossil

---

223 Shearer et al., supra note 221, at 32–33.
fuels. The only exception was yet another delay in the long-mandated, long-postponed shut-down of water-cooling systems needed to operate older natural gas electric generating plants; these plants continue to provide “base load” reliable electricity supplies to California when the sun does not shine and the wind does not blow. The state has an increasing, but still trivial, amount of battery capacity to meet power needs when intermittent renewables are not generating electricity. Regulatory backlash against renewable energy projects has grown in other states (particularly in opposition to wind), but CEQA lawsuits provide another potent anti-renewable tool for ready use by project opponents.

4. CEQA v. Agriculture

The most noteworthy CEQA lawsuits against agriculture were dubbed “the Pistachio wars” after the state’s largest pistachio processor used CEQA to sue competing pistachio processors. In an unusual but welcome decision, the competitor was denied standing to use CEQA to advance its competitive agenda; an appeal was subsequently filed, then dropped. Multiple lawsuits were also filed against winery projects, and one feedlot.

5. CEQA v. Fun

Entertainment and recreational projects that draw more people to a community are an ever-present category of CEQA lawsuits. This Study Period included lawsuits challenging bungee jumping, a golf course’s ongoing existence owed to its new use of recycled water, and sports stadium gifted with project-specific CEQA Legislation in a longstanding legislative tradition resembling papal indulgences granted to naughty aristocrats in the Middle Ages. This CEQA lawsuit category is particularly challenging for seasonal festivals sponsored by local government and community groups that operate on a shoestring budget; both CEQA compliance costs and litigation defense costs can be

---


231 Ruling on Petition for Writ of Mandate at 5–6, Wonderful Citrus II, LLC v. County of Tulare, 2021 Cal. Super., No. VCU283508 (“Wonderful’s interest in this litigation is as a direct economic competitor with direct commercial and competitive interests adverse to Touchstone’s farming operations, not a party motivated by concerns relating to public rights . . . .”).
permanent festival killers. Festivals were again challenged in CEQA lawsuits filed during the Study Period.

6. CEQA v. Retail

Non-union and discount retailers—from Costco to the corner gas station—were targeted by these lawsuits. CEQA lawsuits can be filed anonymously by never-before-in-existence “ad hoc associations” that do not have to disclose any individual member or funding source.232 Retail CEQA lawsuits have long been associated with unions seeking to block non-unionized big box retailers and economic competitors. In this Study Period, the retail sector is no longer experiencing explosive growth of new stores and is, in many cases, retracting, so these CEQA lawsuits were more likely to challenge additions to existing stores (e.g., gas fueling stations). Small business competitors also use CEQA lawsuits against each other.

7. CEQA v. Hotels

The union representing hotel workers has been much more active during this Study Period. These lawsuits target non-union hotels (and projects that include a hotel but for which no hotel owner/operator has been identified). These typically settle with union hotel worker hiring agreements. While national labor laws preclude most unions (except, for example, construction trades and agricultural workers) from engaging in workforce bargaining tactics with non-employers, this union use of CEQA against non-employers has remained a longstanding staple of CEQA lawsuits. Some hotels include one or more apartments reserved exclusively for occupancy of hotel staff, but in our study methodology, hotels with employee-only housing units were not counted as residential projects.

8. CEQA v. Not Much Else

Not on the list: new manufacturing, mining (lithium valley for batteries), transit, forestry, highways, airports, hospitals, or much of anything else. California has lagged far behind the rest of the country in creating new manufacturing jobs233 (typically higher wage jobs available to workers without college degrees). California’s population has continued to decline: the state lost

---

232 See Hernandez, In the Name of the Environment I: 2010-2012, supra note 22, at 24, 78 (finding that only 13% of CEQA lawsuits filed between 2015 and 2018 were filed by known advocacy groups with a history thereof).
500,000 residents between 2020 and 2022, and San Francisco logs in as having the steepest population drop in homebuyers of any large city in the nation, down to its pre-tech boom 2012 level. Mine applications—for gold and lithium—are deep in the CEQA compliance (pre-permit approval) stage.

C. CEQA v. Non-Residential Agency Plans and Regulations

As shown in Figure 6, public agency approvals of plans and regulations that do not allow or incentivize additional housing, and do not approve physical construction activities for any one project or location, but instead result in foreseeable changes to the environment as these plans and ordinances are implemented, were also targeted in our smallest category of CEQA lawsuits.

---


1. CEQA v. Climate Change

There is one stand-out category of public agency approvals for non-residential plans and regulations, and it is climate. California’s climate change commitments made renewable energy—electricity generated by the sun, wind, biogas, or hydropower—second only to water as the top target of CEQA lawsuits in the public/infrastructure/community project category (Figure 4). California’s climate change plans and regulations—especially those aimed at phasing out or banning continued extraction of oil and banning use of natural gas in new homes, restaurants, and other buildings—are the top targets of regulatory agency plans and activities that result in physical changes to the environment only from subsequent plan implementation and regulatory compliance activities.

California has historically pursued both oil and gas extraction, and various longstanding statutes continue to require state and local agencies to authorize these activities.238 At issue are billions of dollars in oil and gas reserves, most of which are now owned by families or smaller companies as many of the major energy companies have liquidated their California holdings.239 Also at issue are hundreds of thousands of jobs, mostly held by those without college degrees, mostly paying above-median wages and benefits in areas where replacement higher wage jobs are unavailable for comparably skilled workers.240 Government agencies also derive hundreds of millions of dollars in tax revenues from this industry, and from related construction, maintenance, and support services and business.241 Shutting down California’s oil industries, long sought by environmentalist “keep it in the ground” advocates, simply means that more oil will be imported from other countries, most notably Saudi Arabia and Venezuela (and, until recently, Russia), with considerably less regard for environmental protections, worker safety, and the rights of disadvantaged communities.242 An authoritative study by the California Council
of Science and Technology concluded that more GHG emissions would be produced from less environmentally stringent oil production practices overseas, and still more would be emitted in transporting oil thousands of miles across multiple oceans and seas. The historically most productive oil reserves on land are located more than 1500 feet below ground on parched lands used intermittently for grazing, and therefore likely have minimal impacts on groundwater.

California has long suffered from the highest gasoline prices in the nation (excepting on occasion Hawaii, which imports all of its oil), and high gas prices have a regressive impact on lower income workers who are more likely to need to be physically present at work to be paid, and to live in less costly areas and have longer commutes. California’s EPA has concluded that, “[t]he highest levels of diesel PM are near ports, rail yards and freeways,” and, as depicted in the CalEnviroScreen4.0 mapping tool, are contrasted with proximity to stationary industrial

243 See JANE LONG ET AL., AN INDEPENDENT SCIENTIFIC ASSESSMENT OF WELL STIMULATION IN CALIFORNIA VOLUME II 40–41 (Cal. Council on Sci. & Tech. et al. eds., 2015), http://ccst.us/wp-content/uploads/160708-sb4-vol-II-7.pdf [http://perma.cc/A4W6-QPX3] (“Oil produced in California using hydraulic fracturing also emits less greenhouse gas per barrel than the average barrel imported to California. If the oil and gas derived from stimulated reservoirs were no longer available, and demand for oil remained constant, the replacement fuel could have larger greenhouse emissions.”); see also CAL. DEP’T OF CONSERVATION, ENVIRONMENTAL IMPACT REPORT, ANALYSIS OF OIL AND GAS WELL STIMULATION TREATMENTS IN CALIFORNIA 12.2-37, 12.2-67 (2015), http://www.conservation.ca.gov/calgem/Pages/SB4_Final_EIR_TOC.aspx [http://perma.cc/6Q3-63H] (rejecting a hydraulic fracturing ban alternative because it “would create much greater significant and unmitigable (Class I) impacts to greenhouse gas emissions” due to increased oil imports which are not subject to California cap and trade requirements, “resulting in an overall net increase in GHG emissions” compared with the status quo).

244 See generally KERN ECON. DEV. FOUND., supra note 239.


facilities such as factories and refineries. Localized health impacts from oil and gas activities have long been alleged, but recent “citizen science” data from thousands of air quality sensors distributed throughout the state and monitored under the supervision of air quality agencies, has confirmed that ground-level pollution exposures are higher for a key fossil fuel combustion pollutant (diesel particulate matter) nearest ports and freeways—not refineries or oil fields. Local agencies have imposed actual or de facto bans (e.g., with mile-wide “buffer” mandates), on continued oil extraction activities, and environmentalists have sued state and local agencies that continue to allow oil and gas extraction as required by existing state law. Opponents to oil and gas extraction agency actions use CEQA to thwart existing legislative mandates, property rights, and the jobs and revenue expectations of tens of thousands of families, in pursuit of speeding up California’s “just transition” to a future without fossil fuel use. California’s gasoline use declined only 1.3% below 2018 levels.

Banning the use of natural gas in new homes, restaurants, and other structures is also a climate policy priority, but natural gas is the last of the less costly energy supplies available to Californians as California’s electricity prices have soared far higher than other states in recent years to fund renewable energy and retrofit existing electricity infrastructure long-neglected by state and utility leaders. A recent study linking natural gas appliances to adverse health outcomes, like asthma, was disavowed by the organization that sponsored the study, which belatedly acknowledged that the study did not assume or

---

249 See generally id.
252 See generally CAL. PUB. RES. CODE §§ 21000–21177 (West 2023).
253 See Ted Goldberg, Strong California Gas Demand Unlikely to Return, as Even Refineries Go Renewable, KQED (Sept. 24, 2020), http://www.kqed.org/news/11839077/strong-california-gas-demand-unlikely-to-return-as-even-refineries-go-renewable [http://perma.cc/KP6A-BW2B]. That same article reports that, as a result of COVID, demand “fell off a cliff.” Id. But that was as of 2020 and does not account for the increasing use of fuel after COVID-related travel restrictions and limitations were lifted.
estimate any causation between natural gas appliance use and asthma rates. Asthma rates have long been known to be higher in communities of color, but researchers have identified access to healthcare and other factors (not use of gas stoves, heaters, clothes dryers, and water heaters) as key culprits. Older homes do not have adequate electricity systems to allow for a simple replacement of natural gas with electric appliances, and civil rights advocates as well as small businesses—such as restaurants reliant on the availability and use of natural gas appliances—have objected and, in some cases, sued to block local “gas ban” ordinances. There is no current plan in California on how to improve the electric grid to accommodate all-electric homes.

Even as the globally tectonic tactics of climate change and a just transition are debated in Congress, the state Legislature, and among a plethora of experts in academia, government, and the non-governmental organizations and private sectors, CEQA lawsuits against agencies seeking to ban or allow fossil fuel extraction and use in California are by far the most frequently targeted agency regulatory action in our Study Period.

IV. CEQA AND THE RULE OF LAW

CEQA is a statute: it was enacted by the Legislature in 1970, and has been amended by hundreds of subsequent statutes over

---


more than fifty years.\textsuperscript{259} Section 21083 of CEQA directs the Governor’s OPR to adopt “guidelines” that:\textsuperscript{260}

- “include objectives and criteria for the orderly evaluation of projects and the preparation of environmental impact reports and negative declarations in a manner consistent with” CEQA;\textsuperscript{261}
- “specifically include criteria for public agencies to follow in determining whether or not a proposed project may have a ‘significant effect on the environment’”;\textsuperscript{262}
- be reviewed and amended “at least once every two years.”\textsuperscript{263}

Guidelines are required to be adopted in compliance with specified sections of the California Administrative Procedure Act (“APA”).\textsuperscript{264} In 1993, the APA was amended to include legislative findings which, among other provisions, concluded that there had been “an unprecedented growth in . . . regulations” including a “complexity and lack of clarity in many regulations” and accordingly directed procedural and substantive requirements for adopting and amending regulations.\textsuperscript{265}

The APA’s procedural requirements include, for example, a mandatory public notice and comment process, and a mandatory evaluation of the economic consequences of regulations, before a new or amended regulation can be approved.\textsuperscript{266} The APA also includes substantive requirements for new and amended regulations:\textsuperscript{267}

(a) “Necessity” means the record of the rulemaking proceeding demonstrates by substantial evidence the need for a regulation to effectuate the purpose of the statute, court decision, or other provision of law that the regulation implements, interprets, or makes specific, taking into account the totality of the record. For purposes of this standard, evidence includes, but is not limited to, facts, studies, and expert opinion.

(b) “Authority” means the provision of law which permits or obligates the agency to adopt, amend, or repeal a regulation.


\textsuperscript{260} See CAL. PUB. RES. CODE § 21083 (West 2005).

\textsuperscript{261} Id.

\textsuperscript{262} Id.

\textsuperscript{263} Id.

\textsuperscript{264} CAL. GOV’T CODE tit. 2, §§ 11340.1–11340.5 (West 2023).

\textsuperscript{265} CAL. GOV’T CODE tit. 2, § 11340 (West 2023).

\textsuperscript{266} CAL. GOV’T CODE tit. 2, §§ 11346, 11346.2–11346.3 (West 2023).

\textsuperscript{267} CAL. GOV’T CODE tit. 2, § 11349 (West 2001).
(c) “Clarity” means written or displayed so that the meaning of regulations will be easily understood by those persons directly affected by them.

(d) “Consistency” means being in harmony with, and not in conflict with or contradictory to, existing statutes, court decisions, or other provisions of law.

(e) “Reference” means the statute, court decision, or other provision of law which the agency implements, interprets, or makes specific by adopting, amending, or repealing a regulation.

(f) “Nonduplication” means that a regulation does not serve the same purpose as a state or federal statute or another regulation. This standard requires that an agency proposing to amend or adopt a regulation must identify any state or federal statute or regulation which is overlapped or duplicated by the proposed regulation and justify any overlap or duplication. This standard is not intended to prohibit state agencies from printing relevant portions of enabling legislation in regulations when the duplication is necessary to satisfy the clarity standard in paragraph (3) of subdivision (a) of Section 11349.1. This standard is intended to prevent the indiscriminate incorporation of statutory language in a regulation.

Section 15000 of the Guidelines state that they are “binding on all public agencies in California.” The CEQA Guidelines have been held to have the same status as regulations. The California Supreme Court has affirmed that, “[a]t a minimum . . . courts should afford great weight to the Guidelines except when a provision is clearly unauthorized or erroneous under CEQA.”

Collectively, CEQA and the Guidelines comprise “The Rule of Law” governing how state and local agencies are supposed to disclose, evaluate, and minimize the significant adverse environmental impacts of discretionary agency decisions to undertake, fund or approve projects, plans, regulations.

A. Administrative Law Jurisprudence v. CEQA Jurisprudence

Under long established principles governing how courts should interpret and enforce statutes and regulations, ordinary administrative law practice is for courts to use an orderly set of “rules” or “canons” to properly interpret and apply the law to

269 See, e.g., Union of Med. Marijuana Patients, Inc. v. City of San Diego, 446 P.3d 317, 323 (Cal. 2019) (“CEQA is implemented by an extensive series of administrative regulations promulgated by the Secretary of the Natural Resources Agency, ordinarily referred to as the ‘CEQA Guidelines.’”).
270 Laurel Heights Improvement Ass'n v. Regents of Univ. of Cal., 764 P.2d 278, 282 n.2 (Cal. 1988) (citing Rural Landowners Ass'n v. City Council, 143 Cal. App. 3d 1013, 1022 (Ct. App. 1983)).
particular situations in dispute. To illustrate both ordinary administrative law jurisprudence and the distinctly different direction CEQA judicial jurisprudence has taken, we will review a 2023 appellate court decision that concluded, for the first time in CEQA’s history, that the noise of future student occupants of an unbuilt dormitory on campus property is an “environmental impact” that was improperly excluded from the EIR prepared by the University of California. Residential neighbors near the campus had produced noise studies confirming the occurrence of late-night student noise, and the record also showed that the campus adopted rules against late night student parties and the campus, city police, and other officials implemented measures (including enforcement of noise ordinance restrictions) to address excessive student noise. The record also included extensive evidence of the absence of proximate student housing for students, high rates of college student homelessness, and the fact that the unavailability of proximate, affordable student housing caused higher student drop-out rates and poorer educational outcomes, especially to students of color and first generation college student.

1. Student Noise and Conventional Administrative Law Practice

In CEQA’s sixty-three year history, there has never been a statute or guideline requiring a noise evaluation of human occupancy of future housing. Some housing—especially housing suitable for families—is more likely to have late noise from colicky-infants, ebullient children playing (and sometimes shouting) during long summer nights, and teenagers fond of loud music even before matriculating to college. In the ordinary administrative law course, a court could not have reasonably concluded that “social noise” from future undergraduate behavior at future dorms was a CEQA impact, using just the most basic canons of administrative law jurisprudence.

272 See Make UC a Good Neighbor v. Regents of Univ. of Cal., 304 Cal. Rptr. 3d 834, 857 (Cal. Ct. App 2023).
273 Id. at 858–59.
a. Plain Language Rule

Section 21083.1 states “that courts not ‘impose’ any ‘substantive requirements beyond those explicitly stated’ provides ‘plain language’ that the Legislature did not intend for the court to expand CEQA to cover social noise...” from future occupants of future housing.275

b. Deference to Expert Administrative Agency Interpretation

OPR, charged with developing the CEQA Guidelines, is the expert CEQA agency in California. The CEQA Guidelines underwent a comprehensive revision in 2018, which, among other features, addressed noise impacts. In Appendix G, Section XI.d, the Guidelines note that “[a] substantial temporary or periodic increase in ambient noise levels in the project vicinity above levels existing without the project” could result in a significant adverse noise impact under CEQA.276 OPR plays a key role in CEQA’s statutory scheme, but it is nevertheless constrained by the statute and, therefore, cannot create a requirement that does not exist in the statute.277 The CEQA Guidelines did not, however, make a shouting college student or illegal late-night loud party a CEQA impact. “Social noise”—if excessive—violates local noise ordinances in public and is a law enforcement issue, not a CEQA impact issue.

c. Deference to Lead Agency Analytical Methodology and Factual Findings

The University’s EIR disclosed the fact that undergraduates were sometimes too noisy late at night and explained what the University was doing—through dorm rules and town-gown policing—to address this unlawful behavior.278 The University’s EIR also evaluated noise impacts from construction, and from post-construction operation (e.g., of building equipment).279

277 See CAL. GOV’T CODE § 11349(b) (West 2001) (providing that all regulations must be authorized by the provision of law that “permits or obligates the agency to adopt, amend, or repeal a regulation”).
278 See Make UC a Good Neighbor, 304 Cal. Rptr. 3d at 858.
279 Id.
d. Consistency with Other Statutes, and with Constitutional Protections

Statutes should be construed to avoid questionable constitutional outcomes, such as differentially assessing the demographics of planned new housing and then speculating as to “social noise” impacts attributable to different ages and races.\(^{280}\)

2. CEQA Jurisprudential Deviations from Administrative Law Norms: Undergraduate Student Dorm Occupancy Example

In January of 2023, an appellate court decided for the first time in CEQA’s history that “social noise”—the noise that individual student occupants of dorms make in the neighborhood immediately north of the UC Berkeley campus—was a CEQA impact.\(^{281}\) “Noise” is indeed a CEQA impact, as identified in the CEQA statute as described by the court.\(^{282}\) The court then turned exclusively to judicial precedent to determine whether “social noise” from future student occupants of unbuilt dorms was also a CEQA impact, fully bypassing conventional administrative law jurisprudential canons in deciding whether this previously unadjudicated issue of unamplified human noise (and specifically unlawful noise from late-night shouts by partying students) was required to be evaluated, and mitigated, under CEQA.\(^{283}\)

Following a familiar pattern of expansive judicial interpretations of CEQA, the court first cited to the California Supreme Court’s first CEQA decision, in 1972, directing that CEQA is to be interpreted by the courts so “as to afford the fullest possible protection to the environment within the reasonable scope of the statutory language.”\(^{284}\) At issue in *Mammoth* was a 184-unit condo project with a restaurant near the Mammoth Mountain ski resort in Mono County.\(^{285}\)

The court went on to acknowledge appellate court decisions, holding that amplified music at a wedding venue in the Santa Cruz mountains, traffic and operational noise from a mining project in the Sierra foothills, and oil well drilling in Kern County were acknowledged to cause noise impacts under CEQA that were insufficiently considered in the CEQA documents (including negative declarations finding the absence of excessive noise) for

\(^{280}\) See, e.g., *People v. Gutierrez*, 324 P.3d 421, 435 (Cal. 2014).

\(^{281}\) See *Make UC a Good Neighbor*, 304 Cal. Rptr. 3d at 857.

\(^{282}\) Id.

\(^{283}\) See *id.* at 857–58.


\(^{285}\) See *id.* at 1052.
those projects. Clearly aware of the controversy the decision would create, the court referred to undisputed studies submitted by neighbor opponents to student dorms that undergraduates were sometimes noisy late at night and that town and gown police enforcement of noise and party restrictions had not eliminated this behavior.

The court then determined that the University improperly concluded that student noise was a behavioral challenge addressed through dorm rules and policing, invalidated the EIR, kept in place an indefinite stay against allowing construction of the dorm project, and remanded the dispute to the lower court to fashion a more specific remedy about how the legally non-compliant EIR must be modified to address this new “social noise” CEQA impact.

3. Next Steps with Social Noise and CEQA

If not accepted for review, and then fully overturned in pending Supreme Court petitions, “social noise” will henceforth be added to CEQA based solely on this new UC Berkeley appellate court dorm decision. Like other CEQA judicial expansions created over the past five decades, this new CEQA impact becomes law without any authorizing action by any elected or appointed state or local officials within or outside the CEQA context.

The appellate court, using CEQA’s expansive tradition of jurisprudence instead of ordinary administrative law canons, used the statutory inclusion of “noise” in CEQA to mandate a new sub-type of unamplified, illegal, late night undergraduate student occupancy “noise.” The court recognized that the plain language of the statute includes “noise,” but then gave no deference to expert agency interpretations in the CEQA Guidelines or a fifty year history of CEQA practice, which collectively never elevated human occupancy “social noise” into a

286 See Keep Our Mountains Quiet v. County of Santa Clara, 187 Cal. Rptr. 3d 96, 111–14 (Ct. App. 2015) (analyzing crowd noise at wedding venue); see also Oro Fino Gold Mining Corp. v. County of El Dorado, 274 Cal. Rptr. 720, 725–26 (Ct. App. 1990) (analyzing noise from mining project); see also King & Gardiner Farms, LLC v. County of Kern, 259 Cal. Rptr. 3d 109, 174 (Ct. App. 2020) (analyzing noise from oil well drilling).
287 See id. at 858, 865. The Court also held that the University failed to justify prioritizing dormitory construction and construction of a homeless shelter at one but not another of the locations identified by the University as suitable for future campus housing. Id. at 863. This portion of the decision is not pertinent to the illustrative example described above.
289 See Make UC a Good Neighbor, 304 Cal. Rptr. 3d at 862–63.
CEQA impact. 291 The court also recognized that the late night student noise at issue was illegal and subject to both university and police enforcement consequences, but concluded that such enforcement had been ineffective in the past in preventing student social noise, so an unknown additional increment of “mitigation” was required by CEQA to prevent the presumptively ongoing but illegal human activities. 292

Unless accepted and overturned by the California Supreme Court, this appellate court decision creates a statewide expansion to CEQA. Because neither the Legislature nor OPR required that the “social noise” of future unbuilt housing be considered an “environmental” impact under CEQA, and there was no legislative debate or APA-compliant public notice and comment process, there is no extant methodology for assessing when, and how, to evaluate and “mitigate” for the “social noise” of future housing occupants in dorms or otherwise. Using conventional CEQA compliance patterns, CEQA practitioners would respond by inventing and unpredictably requiring, for unpredictable agencies and unpredictably for various projects:

- a methodology that includes a demographic prediction of new housing occupants;
- technical methodologies for evaluating the noise of new housing (likely from commissioning studies of “baseline” conditions of noise in occupied housing, such as colicky infants, children hooting when playing tag or hide-and-seek, teenagers playing music in their bedrooms, families using outdoor picnic and play areas, and the baseline and differential frequency of ambulance visits to homes with older versus younger occupants);
- a policy judgment for determining the extent to which noise from housing occupants is “significant” under CEQA, even if otherwise lawful;
- requiring housing projects to include “all feasible mitigation measures” or alternatives to avoid such significant social noise impacts, even if that means that housing should not occur next to noise-sensitive single family homeowners seeking no change in existing ambient noise in “their” environment.

291 See id. at 857–61.
292 See id. at 858, 861.
This is not an exaggerated prediction: noise is just part of the cacophony of objections raised to block new housing—in the name of the environment—to existing neighborhoods within the CEQA framework. For example, in one of the several video documentaries produced by The Two Hundred, “A California for Everyone,” a criminal defense lawyer in downtown Redwood City, who described himself as the “Darth Vader” in his anti-housing zealotry to protect his converted single-family home office in the downtown heart of Silicon Valley, asserted that future occupants of a Habitat for Humanity affordable housing project would cause adverse noise—as well as public safety harms from, for example, leaving tricycles parked on sidewalks. The NIMBY lawyer claimed he had no problem with “those people” [future housing residents] because he represented them in criminal cases. Equating Habitat for Humanity residents with criminal defendants is just one of many examples of the underlying racial bias that makes anti-housing CEQA lawsuits a particular challenge in wealthier and Whiter communities, and importing the demographics of new housing occupants into CEQA invites a wealth of other “new” impact arguments (More crime! More loitering! More home-based car repairs!).

B. CEQA and The Rule of Law

Legislators, major media, and popular opinion have already risen against the use of CEQA by California’s cranky homeowners to block student enrollment in 2022 and to legislatively reverse course on the court’s latest social noise CEQA expansion. Short of full statutory exemptions, however, courts have not confined their discretionary authority to decide CEQA cases to the authority they have been granted by the Legislature.

For example, the Third District Court of Appeal (sitting in Sacramento) heard a challenge to the construction of a replacement office building addition to the historic capitol in Sacramento. An office addition that is decades old was slated

294 Id.
295 Id.
297 See Save Our Capitol! v. Dep’t of Gen. Servs., 303 Cal. Rptr. 3d 761, 771 (Ct. App. 2023); see also Jennifer L. Hernandez & William E. Sterling, California Court of Appeal
to be replaced by a sleek new glass structure—with both bolted onto the historic capitol building.\(^{298}\) The court clearly found aesthetics of the new office building objectionable, since only projects with adverse aesthetics impacts require more analysis under CEQA.\(^{299}\) The Legislature knew their new office project could be sued under CEQA, and they adopted a “buddy statute”—reserved for politically favored projects—that not only streamlined the CEQA judicial review scheduled, but also expressly forbade the court from requiring that the new office project approval be rescinded or that construction be otherwise halted, unless the court found the new office structure caused a significant adverse health and safety impact, or impact to a previously-unknown tribal resource.\(^{300}\) The appellate court ignored the Legislature’s remedy restriction, finding that the Draft EIR failed to adequately depict the new office building and further held that this disclosure failure required full and indefinite cessation of building construction even though only aesthetic and historic (but not health, safety or tribal) CEQA deficiencies were at issue.\(^{301}\)

How did courts decide they could ignore the plain language of CEQA statutes, sidestepping administrative law jurisprudence and the Rule of Law, and instead make a policy choice that favored “the environment” over other policy priorities (coupled with an entirely unbounded definition of what “the environment” actually is)? We will examine one more case illustration before we turn to the Rule of Law discussion and the concluding recommendations.

The community of Encinitas, in northern San Diego County, boasts an average home price of $1.32 million\(^{302}\)—a decrease of 19% from its pre-inflationary high. Quail Botanical Gardens (later transferred to a different operator) operated a visitor center and botanical garden, and sued Encinitas under CEQA for approving the construction of forty new single family homes on a 12.6 acre

\(^{298}\) See Save Our Capitol!, 303 Cal. Rptr. 3d at 773.

\(^{299}\) See id.

\(^{300}\) See CAL. PUB. RES. CODE § 21189.53(a) (West 2018); see also Save Our Capitol!, 303 Cal. Rptr. 3d at 771–72.

\(^{301}\) See Save Our Capitol!, 303 Cal. Rptr. 3d at 805.

lot located in its garden. 303 At issue was the potential that the homes could obstruct ocean views from the ‘garden’s parking lot. 304 The court found that there was no view obstruction impact for adults, but “noted the following:

For a child or disabled person in a wheelchair with a line of vision under a height of four feet, such a limitation would result in total obstruction of certain views of the ocean, leaving, at best, limited and amorphous “view corridors” which were not adequately identified or proven, either quantitatively or qualitatively, during the hearings [by the City to consider approval of the 40-home project]. 305

It is important to note that this jurisprudential pattern of expansive, unpredictable CEQA decisions—including those that are directly at odds with the plain language remedy restrictions of the Legislature—are not partisan. This is not surprising: most Californians, across party lines, strongly support protecting the state’s astounding environment—a popular preference aligned with creatively expanding the scope of “the environment” to be protected by CEQA. 306

1. Off the Rails: Leading Court Cases Creating Modern CEQA Jurisprudence

CEQA was enacted by a nearly-unanimous, bi-partisan California Legislature in 1970 and signed into law by Governor Reagan. 307 CEQA was modelled closely on the National Environmental Quality Act (“NEPA”), enacted a year earlier and signed into law by President Nixon. 308 Both statutes were intended to be applied to projects that were directly undertaken by federal (for NEPA) and state or local (for CEQA agencies), and

304 See id. at 476.
305 See id.
they were not intended to be applied to private projects approved by public agencies. 309 Both statutes were conceived of as largely procedural mandates to evaluate, disclose, receive public input, and then thoughtfully proceed (or not) with a proposed project with full knowledge of the adverse environmental consequences the project would be expected to cause. 310 Both were enacted at a time of bi-partisan consensus—in California following the issuance of an “Environmental Bill of Rights”—that pollution was an acute problem (a river caught fire, a major swath of the California coastline was coated by an oil leak, and choking smog blanketed much of coastal and inland California—including San Francisco Bay and the Central Valley). 311 Beautiful natural places were at risk of irreversible damage (Sequoia National Park was slated to become a Disney resort, much of San Francisco Bay was to be filled to accommodate bulging new Bayfront communities, and San Francisco was to be cleaved by a multi-lane freeway bisecting the city and Golden Gate Park in half). 312 Nature itself (nearly extinct animal and fish species, rare plants, “undergrounded” former streams, and old growth forests) were all at risk of a morning bulldozer assault. 313

For the first few years, CEQA and NEPA continued to track—as did about twenty “baby NEPAs” adopted in various forms by other states. 314 Over the course of the next fifty-two years, NEPA and CEQA sharply diverged. NEPA remained a largely procedural statute, imposing mandates that agencies analyze, disclose, consider feedback, and then explain why they are undertaking an action or issuing a project approval that would cause significant adverse environmental impacts. 315 Persistent efforts, especially in the Ninth Circuit (including California and other Western states), to convert NEPA into a substantive mandate to avoid and minimize significant adverse impacts whenever feasible, repeatedly failed to gain traction with the federal judiciary generally and the Supreme Court particularly. 316

CEQA followed a different pathway, with just a handful of the key early judicial decisions that, in my experience, most shape current CEQA litigation practice and judicial outcomes, as

309 See Hernandez & DeHerrera, supra note 308.
310 See id. at 4, 6.
311 See id.
312 See id.
313 See Lipper, supra note 307.
314 See Hernandez & DeHerrera, supra note 308, at 4–6.
315 See id. at 1–4.
316 See id.
noted below in Table 2. It is noteworthy this expansive judicial interpretation of CEQA was launched with the first Supreme Court decision interpreting CEQA, which applied CEQA to a local agency approval of only 134 condominiums at the Mammoth Mountain ski resort.317

Each of these cases fundamentally changed how CEQA works in practice; many involved judicial elevations of CEQA over other statutes and regulations, especially those directing the approval of more housing. None of these judicial changes to CEQA were required based on a plain language interpretation of any statute or regulation; none of these CEQA expansions were informed by any public notice or comment process; and each was decided in a CEQA-only legal silo that fully ignored other legal imperatives, such as civil rights, housing, and transportation laws. The CEQA directives included in each of these cases was subsequently interpreted and applied by project opponents, along with public agency staff, consultants and lawyers defending the adequacy of CEQA compliance, who collectively invented and at unpredictable intervals through subsequent caselaw modified these judicial decisions.318 Such CEQA directives were interpreted using analytical methodologies, significance criteria, and mitigation measures, by a collection of CEQA practitioners in the public sector, private sector consultants, and private sector attorneys. Stakeholder interests not represented by these individual CEQA practitioners—such as the civil rights community focused on housing and jobs—were largely ignored in a CEQA-only practitioner silo that continued to attempt to prepare legally sufficient CEQA documentation with ever-evolving CEQA deficiency decisions by the more than fifty reported appellate court cases decided each year.

318 The state’s expert CEQA agency, the Office of Planning and Research, was directed by the Legislature to regularly update the CEQA Guidelines, which largely serve as regulations implementing CEQA but are not constraints on further judicial CEQA expansion. See CAL. PUB. RES. CODE § 210083. CEQA Guidelines are infrequently, and only selectively, revised and do not encompass all applicable CEQA requirements as directed in evolving judicial decisions.
TABLE 2: CEQA'S FOUNDATIONAL JURISPRUDENCE

<table>
<thead>
<tr>
<th>Year</th>
<th>Case</th>
<th>Citation</th>
<th>Summary</th>
</tr>
</thead>
</table>
| 1972 | Friends of Mammoth v. Board of Supervisors | 502 P.2d 1049 (Cal. 1972) | CEQA's applicability is expanded exponentially to state and local agency approvals of private project applications, not just projects undertaken by public agencies. The Supreme Court directs judges to use a uniquely broad judicial interpretive rule: CEQA is to be interpreted by the courts so “as to afford the fullest possible protection to the environment within the reasonable scope of the statutory language.”

This first CEQA Supreme Court decision involved a 184-unit condo project, with a restaurant, near the Mammoth Mountain ski resort in Mono County.

This first CEQA Supreme Court decision stands in stark contrast to longstanding “canons” of judicial interpretation and enforcement of statutes, which direct the courts to weigh various factors such as legislative intent, consistency with other laws, and textual clarity or ambiguity.

Since 1970, virtually all modern environmental laws were subsequently enacted, ranging from federal and California versions of the Clean Air and Clean Water Acts, Coastal and Desert Protection Acts, Historic Preservation and Tribal Resource Protection laws, hazardous waste and hazardous materials laws, worker and public health protection laws, earthquake and flood protection laws, wildfire prevention and protection laws, endangered and rare species and habitat protection laws, climate change laws, and scores of sustainable resource protection laws covering groundwater, public and private surface lands blanketing the entire state (except for tribal lands which remain largely under tribal sovereign control), and waters and wetlands.

The First Appellate District, which ruled against the Regents of the University of California and, for the first time in history, concluded that CEQA required analysis and mitigation measures for illegal “social noise” from undergraduate parties, called out this 1972 quote as the “foremost principle” of CEQA and attributed it to the Legislature’s intent, as reported in the 1972 Supreme Court decision. In fact, the Legislature has not enacted this “foremost principle” as the “Legislature’s intent” in

320 See id. at 1052.
321 See BRANNON, supra note 271, at 20.
323 See Make UC a Good Neighbor v. Regents of Univ. of Cal., 304 Cal. Rptr. 3d 834, 843 (Ct. App. 2023).
CEQA; instead, CEQA includes several different statutes reciting the enacted intent of the Legislature, which, among other provisions, notes that CEQA is intended to “[e]nsure that the long-term protection of the environment, consistent with the provision of a decent home and suitable living environment for every Californian, shall be the guiding criterion in public decisions.”

These and other enacted statements of legislative intent under CEQA are rarely quoted, or invoked, in judicial decisions interpreting CEQA.

In 1972, CEQA was a short, general statute, which was of course entirely uninformed by all subsequently-enacted environmental and public health protection statutes. The California Supreme Court did, however, expressly recognize that CEQA was simply a statute: the Court’s direction was that CEQA be broadly construed “within the reasonable scope of the statutory language.”

The 1972 decision has subsequently been relied upon to ignore administrative law jurisprudence. Statutory construction is of more than historical relevance, even as the Legislature has periodically attempted to reign in expansive court interpretations of CEQA with new statutory provisions that continue to be generally, and even expressly, ignored by courts—as discussed below.

---

**1974**

*No Oil, Inc. v. City of Los Angeles, 529 P.2d 66 (Cal. 1974)*

An unprecedented new “fair argument” standard of review is established by the California Supreme Court for the less costly, streamlined “Negative Declaration” environmental compliance created by the Legislature in CEQA for projects that have no or negligible adverse impacts on the environment. The Supreme Court held that a full EIR, which in practice cannot be completed in less than ten months and often takes two years or longer, and cannot be completed for less than $300,000, often inclusive of technical reports and costs in excess of $1,000,000, is required when a project opponent argues that there is a “fair argument” that there “may” be a single significant adverse environmental impact from a project. EIRs remain subject to the “substantial evidence” standard of review, and for practitioners over the next couple of decades, unless an agency is caught in a lie or is openly defiant of a mandatory EIR component like the need to study a reasonable range of alternatives, EIRs are overwhelmingly likely to survive CEQA litigation challenges. CEQA practice evolved into doing an EIR if the project was likely to be sued by someone with money or other resources, even if the project was environmentally benign or beneficial.

---

324 [CAL. PUB. RES. CODE § 21001(d)](http://example.com) (emphasis added).
325 [See lipper, supra note 307.](http://example.com)
327 [See No Oil, Inc. v. City of Los Angeles, 529 P.2d 66, 75 (Cal. 1974).](http://example.com)
328 [See id. at 74–75.](http://example.com)
<table>
<thead>
<tr>
<th>Year</th>
<th>Case</th>
<th>Federal/Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974</td>
<td>San Francisco Ecology Center v. City and County of San Francisco, 122 Cal. Rptr. 100 (1975)</td>
<td>CEQA took another sharp deviation from NEPA when the First District Court of Appeal held that CEQA requires that environmental protection be elevated to a “paramount” and urgent concern, “requir[ing] decision-makers to assign greater priorities to environmental values than to economic needs.” It is no longer enough to analyze, disclose, receive input, and have to explain why an agency is approving a project that will harm the environment. This and subsequent cases held that agencies may not approve a project unless they first require all “feasible” means of avoiding or minimizing significant adverse impacts, while achieving all or most of the project objectives, through a combination of “mitigation measures” aimed at reducing impacts and “alternative” modified projects and/or project locations.</td>
</tr>
<tr>
<td>1990; 2009</td>
<td>Land Waste Management v. Contra Costa County Board of Supervisors, 271 Cal. Rptr. 909 (1990); Schellinger Brothers v. City of Sebastopol, 102 Cal. Rptr. 3d 394 (2009)</td>
<td>The Legislature enacted the Housing Accountability Act (“HAA”) in 1982, which it supported eight years later with 1990 amendments with formal legislative findings that noted that “California housing has become the most expensive in the nation,” a circumstance “partially caused by activities and policies of many local governments which limit the approval of affordable housing, increase the cost of land for affordable housing, and require that high fees and exactions be paid by producers of potentially affordable housing,” and recognized that “[t]he lack of affordable housing is a critical problem which threatens the economic, environmental, and social quality of life in California.” After the Legislature’s housing production increase bills from the 1980’s, including acknowledgement that California’s housing supply was not keeping up with its population growth, state housing costs continued to spiral well ahead of national housing costs for the next four decades. The national housing costs are now far beyond levels affordable to hard working California families. The national average is that median priced homes cost about 4.5 times more than the</td>
</tr>
</tbody>
</table>

329 S.F. Ecology Ctr. v. City & County of San Francisco, 122 Cal. Rptr. 100, 103–04 (1975).
330 See id. at 106–07, 109 n.8 (citing CAL. PUB. RES. CODE § 21100).
333 See id.
In California, homes cost 9 times more than the annual median income. The original version of the HAA required cities and counties to approve housing projects that complied with applicable General Plan and zoning requirements. At the same time (and even before) this early 80s-era housing emergency, the Legislature decided that local governments were taking too long to review and approve development projects that complied with local General Plan and zoning requirements, and imposed a strict schedule for completing the application and approval process in the Permit Streamlining Act (“PSA”). The same year it enacted the HAA, the Legislature enacted a “deemed approved” remedy if an agency missed compliance deadlines, which allowed the applicant to proceed with construction even if a local permit was not issued. Development critics objected to both the HAA and PSA, and, in a political compromise, the Legislature decided not to amend CEQA to conform to these new HAA and PSA mandates. Courts thereafter concluded that the PSA’s “deemed approved” mechanism could not bypass CEQA compliance. Courts also concluded that the time deadlines imposed under the PSA and provided under CEQA, were not effectively enforceable in court. Courts also declined to enforce the HAA in situations where a local agency had yet to certify an EIR pursuant to CEQA. CEQA’s ardent environmentalist supporters, in anti-housing strongholds like Marin County, had their clear first triumph as the “law that swallowed” housing law in California.

1987

*Friends of Westwood, Inc. v. City of Los Angeles*, 235 Cal. Rptr. 788 (Ct. App. 1987)

In practice and in most, but not all, local jurisdictions (San Francisco being the most noteworthy exception), CEQA was not generally applied in cities to private construction projects (e.g., for residential and commercial uses) if the project complied with local General Plans, local zoning, and other code requirements; these

---

334 See Home Price to Income Ratio (US & UK), supra note 332.
335 Id.
340 See Land Waste Mgmt., 271 Cal. Rptr. at 916.
341 See, e.g., id.; Schellinginger Brothers v. City of Sebastopol, 102 Cal. Rptr. 3d 394, 404 (Ct. App. 2009).
342 See Schellinginger Brothers, 102 Cal. Rptr. 3d at 405–06.
343 See id.; Nieves, supra note 21.
projects were considered “by right” and entitled to receive an approval. As land use planning practice evolved, however, local governments began requiring “conditional use permits” ("CUPs") for more categories of projects, notably including apartment projects. A CUP process requires a city to notify and consider input from the public, and also allows a city to impose discretionary conditions of approval on a project, such as specifying the location of a driveway in relation to a street when there was no express or objective zoning standard governing the driveway locations. The City of Los Angeles approved one of the first high rise multi-family housing projects—on Wilshire Boulevard near UCLA—with a CUP that included a few pages of “conditions” the project was required to meet.

In *Friends of Westwood, Inc. v. City of Los Angeles*, the Second Appellate District held that the CUP process was a fully “discretionary” decision by the city, did trigger CEQA compliance, and ordered project approvals rescinded pending CEQA compliance. *Friends of Westwood* set the template for CEQA’s applicability to locally authorized housing projects, which are consistent with General Plan and zoning requirements, but are nevertheless first required to complete the CEQA process. Once an agency concludes that a project will result in a “significant impact to the environment,” CEQA authorizes the agency to deny the project application even for projects that comply with the General Plan and zoning requirements.

1987 was the second big anti-housing “win” for CEQA’s status quo defenders, subjecting even fully compliant housing to extensive study delays and excess costs as each new apartment project (among other housing types) was required to do its own CEQA studies, including studies of “cumulative impacts” that were theoretically supposed to be consistent across a jurisdiction. In my experience, CEQA practice in a small but wealthy city (San Francisco, with forty-nine-square miles of entrenched NIMBYs), began to deviate massively from less wealthy cities where population (and housing) was still increasing significantly—especially in the thousands of square miles comprising Los Angeles, Orange, Riverside and San Bernardino counties.

---

345 *See Friends of Westwood, Inc. v. City of Los Angeles*, 235 Cal. Rptr. 788, 790 (Ct. App. 1987).
346 *See id.* at 800–01, 803–04.
348 *Id.*
Any theoretical understanding that CEQA was a state law that applied in a mostly uniform manner to the same kind of project (apartment building) statewide simply confirmed the absence of the practitioner’s familiarity with other jurisdictions. I was fortunate to be the first CEQA attorney to work in-house for the University of California on all of its campus and hospital projects from 1986-1989. While UC campuses tended to be located in wealthier communities, and town-gown conflicts had already been metastasized into CEQA lawsuits near the oldest campuses near the wealthiest neighborhoods, my legal job provided a vivid education in just how differently a hot-button issue (e.g., traffic congestion) was studied and mitigated (or not) under this supposedly uniform state law.

1988

**Citizens of Goleta Valley v. Board of Supervisors of Santa Barbara, 801 P.2d 1161 (Cal. 1990)**

CEQA had previously been held to require that public agencies consider alternative locations for proposed projects which could avoid one or more significant adverse impacts.\(^\text{350}\) In *Goleta*, the California Supreme Court held that private owners must also consider alternative locations for a proposed project (at issue was a waterfront hotel in Santa Barbara), even if the applicant did not own or control any other site.\(^\text{351}\) Because of this new rule, *Goleta* spawned a cottage industry of specialists who would comb through real estate listings, find potentially suitable sites, document whether or not they were available for purchase, and then either consider them as alternative sites or conclude that no alternative sites were available.\(^\text{352}\) For CEQA practitioners, *Goleta*, like all CEQA published court decisions, had an immediate and retroactive effect in that it simply interpreted existing law. Opponents of the hotel project were of course not mollified by a new EIR, and filed a second CEQA lawsuit – which they lost.\(^\text{353}\) The CEQA compliance process delayed construction for more than a decade.

---

\(^{350}\) See *Citizens of Goleta Valley v. Bd. of Supervisors of Santa Barbara*, 801 P.2d 1161, 1169 (Cal. 1990). The Court affirmed the principle that:

> [A]n EIR for any project subject to CEQA review must consider a reasonable range of alternatives to the project, or to the location of the project, which: (1) offer substantial environmental advantages over the project proposal. . . and (2) may be “feasibly accomplished in a successful manner” considering the economic, environmental, social and technological factors involved.

*Id.* at 1168 (citations and emphasis omitted).

\(^{351}\) *Id.* at 1180.

\(^{352}\) See Lennie Rae Cooke & Craig Stevens, *CEQA Portal Topic Paper: Alternatives*, AEP CEQA Portal 6 (Oct. 18, 2018), http://ceqaportal.org/tp/Alternatives.pdf (http://perma.cc/X6XP-MRHY) (providing that offsite alternatives should be considered); *see also* [CAL. CODE REGS. tit. 14, § 15126.26(a) (2023)](http://perma.cc/X6XP-MRHY) (providing that “[a]n EIR shall describe a range of reasonable alternatives to the project, or to the location of the project, which would feasibly attain most of the basic objectives of the project”) (emphasis added).

The result of the Table 2 cases was to make housing—like other routine construction in ordinary communities built in compliance with California’s ever-more-stringent environmental, building, conservation, labor, and public health standards—subject to CEQA. Judges are lawyers, and most lawyers find the idea of a short delay to obtain a more comprehensive understanding to be a routine and beneficial part of law practice. CEQA also has the beguiling feature of being widely understood to just require “more study”—after all, agencies could still re-approve a project once they “fully complied” with CEQA. The consequences of those harmed by delayed projects, and projects that became more costly or were derailed entirely by delays that coincided with shifting economic and political conditions, are often not part of the “CEQA administrative record” at issue in CEQA lawsuits; even if included in the record, they would likely be considered subordinate or irrelevant to the overarching 1972 California Supreme Court directive that CEQA should be broadly interpreted by the courts to protect the environment.

2. 1993 and Beyond: Legislature’s Largely Failed Attempts to Restore Administrative Law Jurisprudence to CEQA

CEQA in practice is different from the CEQA lawsuit briefs about the parsed merits of any particular sub-argument involving a sub-issue of one of the scores of impact categories that have been sufficiently addressed in the CEQA process. The practical, political, economic, and policy implications of the vast expanse of CEQA through CEQA jurisprudence was accordingly considered in the Legislature, which responded in several rounds, but most broadly in 1993 with statutory amendments to CEQA designed to bring greater predictability to CEQA.\(^\text{354}\)

With continued underproduction of housing, along with high profile CEQA lawsuits against infrastructure and educational projects, an environmental leader and hero in the Legislature, Senator Byron Sher of Palo Alto (also a Stanford Law School professor) led a two-bill, generally bi-partisan effort to reform CEQA in 1993 with a series of statutory changes designed to accelerate the CEQA compliance schedule, reduce compliance costs, and make judicial outcomes more predictable.\(^\text{355}\) In my opinion, this 1993 Legislation was the only year, in fifty years,


\(^{355}\) See id.
where broadly applicable CEQA reforms were not politically killed by CEQA’s most powerful status quo defenders in the State Building and Construction Trades Council (“Building Trades”) and environmental advocacy groups.

For example, the Legislature amended CEQA to require that there must be “substantial evidence” in support of a “fair argument” that a project would have a one or more significant adverse environmental impacts, in an effort (largely unsuccessful) to make negative declarations more defensible. In a fifteen-year study of CEQA lawsuit outcomes, nearly sixty percent of negative declarations failed to withstand judicial scrutiny. Courts concluded that even air conditioner noise was enough to invalidate a negative declaration and require an EIR. Courts also concluded that even non-expert opinion, about noise from trucks, met the “substantial evidence of a fair argument” standard.

Also in 1993, the Legislature attempted to fix the “fit the punishment to the crime” CEQA remedy problem. In law school, and in civil litigation, adequately analyzing and mitigating ninety-eight percent of impacts covered in two hundred pages of EIR text, supported by five hundred pages of technical appendix, could earn an “A” grade and meets all standards of review normally applied by civil courts (preponderance of the evidence, substantial evidence, etc.). Under CEQA jurisprudence, in contrast, a judicial conclusion that the agency fell short of full CEQA compliance for just two percent of the analysis (just a handful of pages comprising subparts of one or two environmental impact topics) most commonly results in the judicial remedy of rescinding all project approvals and re-doing the EIR process to fix the deficient analysis in a process that takes a year or longer. The Legislature’s fix was directing the courts to order “severance” so whatever portion of a project that was not affected by the deficiency could proceed without further delay. One appellate court district steadfastly

356 See CAL. PUB. RES. CODE § 21082.2 (West 2023).
360 See CAL. PUB. RES. CODE § 21168.9 (West 2023).
361 See id. § 21167.1.
declines to authorize any severance remedies,362 and the others do—but trial courts are split and the results are unpredictable. Notwithstanding this severance remedy directive, the direct practical consequence of being targeted by a CEQA lawsuit is being exposed to a potential judicial rescission remedy—enough to dissuade most lenders, investors, and grantors in funding a project while a lawsuit is pending, without regard to the lawsuit’s merits, and without any of the normal safeguards (including bond requirements) of judicially-imposed preliminary injunctions pending the merits decisions.

Section 21005(b) was also added in 1993: “It is the intent of the Legislature that, in undertaking judicial review [in CEQA lawsuits], courts shall continue to follow the established principle that there is no presumption that error is prejudicial.”363 Courts largely declined to give this statute any practical effect, subsequently holding, for example, that a disclosure omission is prejudicial in precluding informed public participation,364 that the burden falls on the lead agency to demonstrate that an error is not prejudicial,365 and most significantly “when an agency fails to proceed’ as required by CEQA, harmless error analysis is inapplicable.”366 One important California Supreme Court decision did conclude that a transit project EIR which failed to analyze any environmental impacts against the present-day “baseline” of existing environmental conditions, and thereby failed to include any meaningful assessment of construction impacts such as noise, dust, and air pollution, was erroneous—but that the public had a common sense understanding of construction impacts, so meaningful public engagement could occur even in the absence of an EIR analysis of these impacts.367

Finally, the Legislature enacted section 21083.1 of CEQA, which reads in full:

It is the intent of the Legislature that courts, consistent with generally accepted rules of statutory interpretation, shall not

---


363 CAL. PUB. RES. CODE § 21005(b).


interpret this division or the state guidelines adopted pursuant to Section 21083 in a manner which imposes procedural or substantive requirements beyond those explicitly stated in this division or in the state guidelines.\textsuperscript{368}

This statute was cited in two Third District Court of Appeal decisions, including one that noted courts are “constrained to reject” interpretations of CEQA that are “beyond the explicit terms of the act”—even if accepting the interpretation would “arguably afford greater protection to the environment.”\textsuperscript{369}

As noted above, in my experience, the 1993 Session was the last time that the Legislature attempted to reform CEQA using traditional statutory amendment tools: new and amended statements of legislative intent, legislative directives regarding the absence of prejudicial error, legislative directives to allow portions of the project not affected by an analytic deficiency to proceed with a severance remedy, and—more importantly—a clear direction that courts no longer construe CEQA “broadly to protect the environment” but instead avoid construing CEQA “in a manner which imposes procedural or substantive requirements beyond those explicitly stated” in the CEQA statute or guidelines.

None of these statutes resulted in any meaningful change to CEQA jurisprudence. CEQA lawsuits are filed in a form of litigation proceeding called a “writ of mandamus”\textsuperscript{370}—an old common law term that differs from ordinary civil disputes initiated by a “complaint.”\textsuperscript{371} Writs seek to compel agencies to undertake, or refrain from undertaking, an action.\textsuperscript{371} Although a writ of mandamus is entirely ordinary in CEQA, in U.S. law it is considered an “extraordinary writ” that “is not [issued as] a matter of right, nor governed entirely by fixed rules, but is within the ‘sound’ or ‘wise’ discretion of the court[s].”\textsuperscript{372}

In contrast to ordinary administrative law jurisprudence, California courts have demonstrated no appetite to have their

\textsuperscript{368} CAL. PUB. RES. CODE § 21083.1 (West 2023).
\textsuperscript{369} Picayune Rancheria of Chukchansi Indians v. Brown, 178 Cal. Rptr. 3d 563, 573 (Ct. App. 2014); see also W. Placer Citizens for an Agric. & Rural Env’t v. County of Placer, 50 Cal. Rptr. 3d 799, 806 (Ct. App. 2006).
\textsuperscript{370} See STEPHEN L. KOSTKA & MICHAEL H. ZISCHKE, PRACTICE UNDER THE CALIFORNIA ENVIRONMENTAL QUALITY ACT § 23.2 (2d ed. 2008).
\textsuperscript{371} See Writ, BLACK’S LAW DICTIONARY (11th ed. 2019) (defining a “writ” as “[a] court’s written order, in the name of a state or other competent legal authority, commanding the addressee to do or refrain from doing some specified act”).
discretion boxed in by statutes or rules in CEQA jurisprudence, as shown in Table 3:

<table>
<thead>
<tr>
<th>Year</th>
<th>Case Title</th>
<th>Court Decision</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>Mountain Lion Foundation v. Fish &amp; Game Commission, 939 P.2d 1280 (Cal. 1997)</td>
<td>Courts rejected newly-enacted statutory constraints on CEQA, again affirming that CEQA prohibits agencies from approving a project causing a significant adverse impact if there are “feasible alternatives or mitigation measures” available to substantially lessen that effect.</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>Berkeley Keep Jets Over the Bay Committee v. Port of Oakland, 111 Cal. Rptr. 2d 598 (Ct. App. 2001)</td>
<td>EIRs, formerly largely defensible in court, become newly vulnerable with an expansive new application of the “prejudicial abuse of discretion” standard (previously used to review an agency’s compliance with CEQA’s procedure) to evaluate the substantive adequacy of an agency’s analysis of impacts. The court found that the Port did use a protocol approved by an expert agency to evaluate toxic air emissions from an airport expansion project. The court further found that the Port knew about and was advised that more recent draft protocols had been developed but not yet adopted, but were, in the opinion of a staff member of the expert agency, the “best available data” and should be used. The Port continued to use the approved protocol, while responding on the record to arguments that the newer draft protocol should be used and while adopting nine mitigation measures to reduce toxic air emission exposures at the airport project. The appellate court concluded that the Port erred in not using the new draft protocol endorsed by staff of the expert agency and invalidated the EIR. This decision reverses prior court precedent of deferring to agency conclusions on factual issues when supported by substantial evidence in the record even in an EIR context, as well as longstanding case law that disagreements among experts are resolved in favor of the CEQA lead agency. This decision launches a new era of challenges to the analytical</td>
<td></td>
</tr>
</tbody>
</table>

373 Mountain Lion Found. v. Fish & Game Comm’n, 939 P.2d 1280, 1298 (Cal. 1997).
374 Berkeley Keep Jets Over the Bay Comm. v. Bd. of Port Comm’rs, 111 Cal. Rptr. 2d 598, 606 (Ct. App. 2001).
375 See id. at 613–15.
376 Id. at 613–15.
377 See id. at 613–14.
378 See id. at 615.
sufficiency of EIRs: courts began to substitute their own judgment for the agency’s factual determinations by an expansive application of the “abuse of discretion” standard formerly applied to procedural violations of CEQA.

<table>
<thead>
<tr>
<th>2002</th>
</tr>
</thead>
</table>

**Communities for a Better Environment v. California Resources Agency, 126 Cal. Rptr. 2d 441 (Ct. App. 2002)**

Amendments to CEQA’s implementing regulations, the CEQA Guidelines that attempted to integrate more than fifty major environmental protection laws (clean air and water, protected species and resources, etc.) enacted since 1970 were rejected by the Third District as inconsistent with the “fair argument” standard of review and other CEQA precedents. When, where, and what CEQA requires in “additional” evaluation and/or mitigation beyond compliance with applicable environmental and public health statutes and regulations remains entirely unpredictable in CEQA litigation. These regulatory amendments to CEQA, which are subject to the APA, followed from several court decisions confirming that compliance with an applicable environmental and/or health protection standard did satisfy the “substantial evidence” standard of review in CEQA for showing that an impact was reduced to a less than significant level. Both Sundstrom and Leonoff were decided under the far less deferential negative declaration “fair argument” standard. However, the Third District rejected amendments to the Guidelines codifying these earlier published judicial decisions, largely based on an expansive reading of the “fair argument” standard of review applicable only to negative declarations (and not environmental impact reports or exemption determinations). It is also noteworthy that the Third District decision completely ignores the “non-duplication” criteria of the APA, which requires that regulations not duplicate other laws or regulations and is itself an ordinary canon of administrative law jurisprudence that warrants the full integration of the CEQA Guidelines with the now thousands of environmental, safety, and health protection laws and regulations that have become effective since 1970.

---


380 See Leonoff v. Monterey Cnty. Bd. of Supervisors, 272 Cal. Rptr. 372, 382 (Ct. App. 1990) (holding that compliance with hazardous materials management laws was sufficient to conclude that hazardous materials impacts are less than significant); Sundstrom v. County of Mendocino, 248 Cal. Rptr. 352, 360 (Ct. App. 1988).

381 See Leonoff, 272 Cal. Rptr. at 382; Sundstrom, 248 Cal. Rptr. at 360.

Express judicial rejection of the 1993 statutory standard (Section 2005(b)) that errors and omissions in CEQA documents is presumed to be non-prejudicial. First the Fourth District, and then most others, held that the “omission” of “important environmental information” is “presumed to be prejudicial error.” Courts differ as to what constitutes “important environmental information” and why. Some courts (including the California Supreme Court) for some projects continue to conclude that the omission of information is not necessarily prejudicial. These inconsistent court conclusions about whether “missing” information or analysis is prejudicial have introduced greater uncertainty to judicial outcomes than existed pre-1993 under the former “substantial evidence” and “fair argument” standards of review. For challenged EIRs, CEQA lawsuits always allege insufficiently-detailed disclosure, analysis, and/or mitigation. The judicial outcome is, as acknowledged by learned University of California environmental law professors, unknowable.

While CEQA political rhetoric often pits “environmentalists” against “developers,” the victims of CEQA jurisprudence are far more likely to be those not served by unbuilt facilities and “the environment” not located immediately adjacent to those forced to live farther away from campus.

California college campus systems, especially the University of California and Cal State University systems, lost a string of CEQA lawsuits based on CEQA mandates newly identified in court decisions decided after the 1993 CEQA reform wave. UC Berkeley CEQA lawsuits are in the news, but anti-campus CEQA lawsuits resulting in blocked campus enrollment growth and development have long been a staple in CEQA jurisprudence, as noted in the following three examples:
Monterey Bay State University must mitigate impacts to local infrastructure and public services due to campus expansion, even if it has no funding to do so. Cost-sharing of infrastructure improvements in mitigation is not rendered infeasible by uncertainty in the local agency’s ability to obtain its matching share of necessary funding.

San Diego State must contribute funds for off-site mitigation of environmental effects of campus expansion, even if the Legislature has declined to appropriate funds to do so. San Diego State must tap other resources, such as alumni, for funding or consider redirecting student enrollment to other campuses.

East Bay (Hayward) State University must analyze and mitigate impacts from increased student use of regional park trails.

“Social Noise” from future undergraduate residents of unbuilt dorms was added to CEQA as an environmental impact.

---

**Center for Biological Diversity v. Department of Fish & Wildlife, 361 P.3d 342 (Cal. 2015)**

Housing projects must consider greenhouse gas emission impacts from new residents in relation to state and global climate science, even though future residents could have a greater impact on greenhouse gas emissions if the challenged housing project is denied and they live somewhere else.

The dissenting Justice opines that CEQA is not a population control statute; his colleagues in this and other opinions agree that CEQA is not a population control statute.

In its recently approved (December 2022) “Scoping Plan” to achieve California’s greenhouse gas reduction targets, CARB reported on an academic study commissioned by CARB and the California EPA to evaluate how CEQA affects housing production. Although the study looked at fewer than twenty

---

387 *City of Marina*, 138 P.3d at 706.

388 *Id.*

389 *City of San Diego v. Bd. of Trs. of California State Univ.,* 352 P.3d 883, 885 (Cal. 2015); *City of San Diego v. Bd. of Trs. of California State Univ.,* 135 Cal. Rptr. 3d 495, 522 (Ct. App. 2011).

390 *City of Hayward*, 195 Cal. Rptr. 3d at 637.

391 *See Make UC A Good Neighbor v. Regents of Univ. of Cal.,* 304 Cal. Rptr. 3d 834 (Ct. App. 2023).

392 *Ctr. for Biological Diversity v. Dep’t of Fish & Wildlife, 361 P.3d 342, 350 (Cal. 2016).*

393 *See id.* at 367 (Chin, J., dissenting).

jurisdictions, in contrast to the comprehensive statewide analysis included in this study, the authors reported that two-thirds of anti-housing CEQA lawsuits claimed an alleged inadequacy of the project’s compliance with GHG provisions of CEQA, and the even newer “Vehicle Miles Traveled” climate metric impact—consisting of estimated future use post-construction of residential automobile/pickup truck use, even by a carpool or electric car. CARB has not created clear, feasible, or lawful standards for how new housing is supposed to mitigate GHG and VMT impacts—an APA violation. For example, a competitor’s 2020 CEQA lawsuit against a veterans outpatient health clinic in Bakersfield alleged that the CEQA documentation prepared by the city insufficiently considered state GHG requirements.

The greatest source of legal uncertainty in more recent judicial opinions derives from increasingly common rejection of the “substantial evidence” standard of review for the analytical environmental content of EIRs. Under the substantial evidence standards, courts defer to lead agency factual determinations as to the appropriate impact assessment methodology, impact significance criteria, and mitigation measure effectiveness when these are supported by substantial evidence in the record. The far less deferential “prejudicial abuse of discretion” standard of review – formerly used mostly to enforce CEQA’s procedural requirements – is now far more commonly applied to judicially reject an agency’s analytical and mitigation determinations. In practice, this means that courts are asked to conclude that an EIR is fatally flawed because the agency did not do an analysis of a particular sub-topic (or sub-topic of a sub-topic) in the Draft EIR itself not simply in response to comments in a Final EIR, and not staff report or hearing responses to “late hit” comments submitted well after CEQA’s public comment periods.

<table>
<thead>
<tr>
<th>Year</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015; 2017; 2018</td>
<td>The greatest source of legal uncertainty in more recent judicial opinions derives from increasingly common rejection of the “substantial evidence” standard of review for the analytical environmental content of EIRs. Under the substantial evidence standards, courts defer to lead agency factual determinations as to the appropriate impact assessment methodology, impact significance criteria, and mitigation measure effectiveness when these are supported by substantial evidence in the record. The far less deferential “prejudicial abuse of discretion” standard of review – formerly used mostly to enforce CEQA’s procedural requirements – is now far more commonly applied to judicially reject an agency’s analytical and mitigation determinations. In practice, this means that courts are asked to conclude that an EIR is fatally flawed because the agency did not do an analysis of a particular sub-topic (or sub-topic of a sub-topic) in the Draft EIR itself not simply in response to comments in a Final EIR, and not staff report or hearing responses to “late hit” comments submitted well after CEQA’s public comment periods.</td>
</tr>
</tbody>
</table>

(providing that of the small percentage of projects studied that were litigated, approximately two-thirds were challenges based on claimed deficiencies in their GHG or VMT analysis).


Examples of this CEQA jurisprudential trend are decisions that an EIR is flawed because it did not include what a court later decides is the best practically available scientific information. For example, the San Diego Association of Governments failed to comply with CEQA by using a methodology containing “data gaps” to estimate the amount of existing farmland (and therefore the project’s impacts to existing farmland), even though the agency explained why its methodology was sufficient and appropriate.398

Similarly, the California Supreme Court concluded that Fresno County committed prejudicial error by not considering in the EIR for a mixed use housing project that the localized toxicity of ambient air pollutants produced primarily from project traffic. Fresno County, as well as two expert state air quality agencies, informed the Court that the impacts of these ambient pollutants that caused regional smog could not be accurately assessed on a localized level. The Court concluded that even if the lead and expert agencies were correct, the EIR was flawed for not explaining the an analysis that was not and could not have been done.399

In 2015, the California Supreme Court concluded that even though impacts to an endangered fish were analyzed and mitigated in an EIR for another mixed use residential project, the EIR was nonetheless flawed because it omitted an analysis of impacts to the juvenile stage this fish. The same Court concluded that the project’s compliance with a statewide target for reducing greenhouse gas emissions was an insufficient CEQA impact significance standard, and identified—but did not endorse—four potential “paths” for completing a legally-sufficient CEQA analysis of greenhouse gas impacts.400

These Table 3 examples are of judicially-created, presumptively mandatory CEQA compliance requirements for which there are no “express” requirements in the CEQA statute or Guidelines requiring analysis or mitigation, and which, accordingly, should not have been found to be prejudicial error gaps under the plain language of Section 21083.1 of CEQA. Absent judicial enforcement of Section 21083.1, CEQA practitioners and agencies working on CEQA documents—particularly those involving well-funded and entrenched project opponents—are routinely slammed with scores of “studies” purporting to show some CEQA impact or another, each hoping

399 See Sierra Club v. County of Fresno, 431 P.3d 1151, 1169 (Cal. 2018).
400 See Ctr. for Biological Diversity v. Dept. of Fish & Wildlife, 361 P.3d 342, 356–57 (Cal. 2015).
that a judge (or group of appellate justices) will conclude that even the most elaborate and costly EIR has a fatal substantive analytical flaw.

These “best scientific data” open-ended judicial precedents impose a vastly uncertain CEQA compliance obligation on agencies, without Legislative or APA-compliant regulatory authority. For example, a recent study of greenhouse gas emissions and climate change reported its review of “88125 climate-related papers published since 2012.”401 The study searched “the Web of Science [online database] for English language ‘articles’ added between the dates of 2012 and November 2020 with the keywords ‘climate change’, ‘global climate change’ and ‘global warming.’”402 The study’s authors found that, over an eight year period, ten thousand scientific articles per year were published on GHG and climate change in English alone.403 No city planner reviewing an apartment project application can sort through and identify the “best available scientific data” in this study tsunami, to accurately guess at what must be included in an EIR.

Greenhouse gas impacts—and global climate change—are CEQA topics especially vulnerable to CEQA lawsuits. For example, in another 2015 case, Friends of Highland Park v. the City of Los Angeles (an unpublished appellate court decision reviewing a twenty-condo, fifty-affordable housing unit project in the Highland Park Transit Village of Los Angeles), the court concluded the project’s greenhouse gas emissions analysis was insufficient and ordered rescission of this small housing project.404

CEQA litigation frequently involves disputes over whether the lead agency used best available scientific data, with courts offering some legal refuge (if an EIR is completed) for studies prepared by a qualified expert, even if other experts disagree. Experts that do not specifically address and rebut the sometimes hundreds of studies lobbed into the lead agency as “comments” on an EIR risk the wrath of a court, however, if opposition studies are not also rebutted by the agency’s expert in the EIR record. This war of experts, on multiple topics, can consume many months and any hundreds of thousands of dollars—all to answer this question:

401 See Mark Lynas et al., Greater Than 99% Consensus on Human Caused Climate Change in the Peer-Reviewed Scientific Literature, 16 ENV’T RSCH. LETTERS 1, 1 (2021).
402 Id. at 2.
403 See id.
does building new homes for Californians, in compliance with the most stringent environmental standards in the world, cause significant adverse climate change impacts? In its recent Scoping Plan, the state’s leading climate agency—the California Air Resources Board—citing to a study prepared by UC Berkeley scholars—acknowledged that greenhouse gas emissions, and the new regulatory climate-based “vehicle “miles traveled” impact, are in dispute in two-thirds of the anti-housing CEQA lawsuits considered in that study.405

3. The Legislature Turns “Transactional” – Favored or Priority Projects Granted Statutory Exemptions from CEQA, Less Politically Powerful Projects Left to Flounder in Uncertainty

The Legislature did not show any further appetite for directing the courts on how CEQA should be interpreted, and instead responded to an ongoing but increasingly notorious practice of enacting more than one hundred statutory exemptions from CEQA, either for specific projects (prisons,406 the 1982 LA Olympics407 in their entirety), or for categories of projects (pipelines in public streets less than one mile long408), the adoption of Groundwater Sustainability Plans,409 and the allocation of new housing planning and approval mandates to cities and counties under the Regional Housing Needs Assessment laws.410 The reach and effectiveness of these exemptions was “transactional” politically: a strong political stakeholder, with support or non-opposition from other political stakeholders, got an unambiguous exemption. Exemptions for the poor (affordable and farmworker housing), the less politically powerful (bike path users), and the destitute (homeless shelters) got political bragging rights but highly restrictive, and time-limited, exemptions.411

405 See CAL. AIR RES. BD., supra note 394, at 19–20; see also MOIRA O’NEILL-HUTSON ET AL., supra note 394, at 5, 83.
406 See CAL. PUB. RES. CODE § 21080.03 (West 2023).
407 See id. § 21080(b)(7).
408 See id. § 21080.21.
410 See CAL. GOV’T CODE § 65584(g) (providing a partial list of statutory exemptions, which can also be found in CAL. CODE REGS. tit. 14 § 15260 et seq.).
411 See CAL. PUB. RES. CODE § 21080.25(b)(1) (exemption for pedestrian bike paths); id. § 21080.27(b)(1) (exemption for City of Los Angeles emergency shelters or supportive housing).
The Legislature also directed the Governor’s OPR to promulgate CEQA Guidelines, adopted as regulations, which are required to identify categories of projects that are exempt from CEQA if they meet all categorical exemption regulatory criteria, and there are no “unusual circumstances” that cause an otherwise environmentally benign project to nevertheless cause a significant adverse environmental impact.412 Construction of a code-compliant single family home on a single family lot is a Class 3 exemption, and in the longest known judicial dispute involving a categorical exemption, one Berkeley home was caught in multiple court proceedings for eleven years.413 There are thirty-three classes of categorical exemptions.414 There is also a “common sense” regulatory exemption for agency actions which could not conceivably result in any change to the physical environment that could be environmentally significant,415 which was originally explained to the author as the need to avoid CEQA for a state agency deciding whether to stock Coke or Pepsi in its vending machines.

The Legislature also enacts non-codified, one-time CEQA exemptions in annual budget trailer bills.416 These are typically enacted in a hurried process to meet budget deadlines when failure to do so means legislators aren’t paid, and they typically involve no CEQA policy committee hearings or other meaningful public disclosure or debate. CEQA compliance can also be avoided if the Governor declares an emergency, albeit with less legal certainty for projects that are approved or funded, but not fully constructed, during an emergency.417

The challenge posed by “transactional” exemptions is that the housing and infrastructure needed by ordinary people does not have the well-funded and well-organized special interest stakeholder sponsors skilled at “making a deal” to avoid CEQA for their particular project or category of projects. Vigorous defense of the CEQA status quo by two of Sacramento’s most powerful constituencies—Building Trades and

412 See id. § 21080(b)(9); see also id. § 21084.
415 Id. § 15061.
environmentalists—also made Legislators wary of touching “third rail” CEQA reforms. Even a once-ardent supporter of CEQA reforms—who led the Senate chamber before becoming the Mayor of Sacramento—settled for a “Kings Arena” buddy bill exemption; the bill was inclusive of remedy restrictions forbidding the court from stopping the project and was introduced and approved in the last two days of the legislative session.418 Other parts of that Legislation, crafted exclusively by environmentalists and Building Tradesthat were heralded as meaningful pro-housing CEQA reforms, were too narrow or otherwise burdensome to have much practical effect in the real world, consistent with the policy objective of these CEQA status quo defenders.

Courts generally uphold CEQA exemptions, especially statutory exemptions. Courts do not, in this context, interpret CEQA expansively to defeat an exemption for a project that would cause environmental harm: the sole legal question is whether the challenged project meets the exemption criteria.419 Categorical exemptions are subject to a less deferential review process and must be “narrowly construed” to effectuate the judiciary’s broad interpretation of CEQA.420

4. Legislature v. CEQA Jurisprudence

CEQA amendments by the Legislature evolved from enacting “transactional” full statutory exemptions from CEQA for specific, politically favored projects that have satisfied environmental, labor and local government stakeholders, to enacting a statutory program for “Environmental Leadership Projects” (“ELP”) that meet eligibility and political stakeholder acceptance criteria as approved by the Governor.421 ELP projects are entitled to “streamlined” judicial review, completing trial and appellate court proceedings in a total of 270 days.422 Few ELP projects are approved, fewer are challenged, and none meet the 270 day

421 See CAL. PUB. RES. CODE §§ 21178–21189.3.
422 See id. § 21185.
deadline, but they do come close based on new Judicial Rules of Court for ELP projects.423

Several of these transactional legislative dispensations expressly limit judicial discretion, forbidding judges from imposing any remedy to stop project construction, or require rescission of project approvals, except under prescribed circumstances.424

The Third Appellate District rejected an express legislative prohibition on CEQA judicial remedy of halting or rescinding a Capitol office building project, unless the project presents an immediate threat to public health and safety, or if the project contains “unforeseen” important cultural or historical artifacts that would be adversely affected by the project’s continuance.425

The Court found that the challenged office project on the state capital had adverse and under-disclosed aesthetic and historic resource impacts and ultimately could not commence construction pending a new and legally compliant EIR process.426

V. RECOMMENDATIONS AND NEXT STEPS

When CEQA was adopted in 1970, there was no Endangered Species Act, Clean Water Act, Clean Air Act, Coastal Protection Act, or any of the myriad new environmental protection statutes initially adopted later in the 1970s, many of which have been strengthened thereafter.

In the void of any meaningful environmental protection mandates except CEQA, the Supreme Court’s 1972 exhortation that CEQA be broadly construed to protect the environment427

---


424 See, e.g., CAL. PUB. RES. CODE § 21168.6(h) (repealed as of inoperative date); see also Save Our Capitol! v. Dep’t of Gen. Servs., 303 Cal. Rptr. 3d 761, 771–72 (Ct. App. 2023).

425 Save Our Capitol!, 303 Cal. Rptr. 3d at 771–72.

426 See id. at 805–06.

was all that stood between an agency determined to authorize old
growth timber clear cuts.\textsuperscript{428} All those projects were stopped, in
part through CEQA, but more meaningfully and permanently
through the dozens of other environmental laws enacted
subsequent to CEQA.

The 1972 directive, though, needs to be revisited to reflect
the reality of CEQA practice today. In Berkeley, repairing our
small kitchen deck was “categorically exempt” from CEQA as a
“repair” of an existing structure. A cranky neighbor could have
sued us and claimed we didn’t qualify for a categorical exemption
based on an “unusual circumstance”—such as the non-compliant
side setback distance to the next-door house (when both houses,
and the broken deck, were built prior to the adoption of side
setback requirements). If we were unwilling to pay the City to
defend us from that lawsuit, and if we were unwilling to
indemnify the City in case our neighbors won and the City was
ordered to pay the neighbor’s attorneys’ fees, then we could not
get approval from the City to repair the deck. Fortunately, we did
not have cranky neighbors. Cranky neighbors love CEQA. Only
wealthy neighbors can pay for CEQA compliance costs, litigation
costs, and fund city indemnity demands. If we had a cranky
neighbor, we would have had to demolish the deck. Bummer, as
it was also our backyard access and fire exit.

This is not CEQA as enacted by the Legislature, nor is it
CEQA as reviewed by the courts. It is CEQA in practice, and the
Legislature (through CEQA statutes), Governor (through the
CEQA Guidelines), and the courts (through CEQA jurisprudence)
should be aware of what CEQA is actually doing, for whom it is
acting on behalf of, and what it is blocking—like housing and
climate resiliency.

CEQA today is about protecting the status quo by stopping
housing, and “those people,” and all the infrastructure “they” need.
CEQA today is about protecting the current “natural”
environment, inclusive of catastrophically mismanaged forests,
crumbling levees, reverse-flow rivers, and water supply shortfalls
that have left one million residents in urban and rural
communities (mostly disadvantaged communities of color) without
water they can safely drink from their taps. CEQA today favors
blocking two-story homes to preserve a parking lot micro-
environment ocean view of four-foot tall children and adults in

\textsuperscript{428} Ebbetts Pass Forest Watch v. Cal. Dep’t of Forestry & Fire Prot., 43 Cal. Rptr. 3d
363 (Ct. App. 2006).
wheelchairs, even as climate change policies demand vast and fast action to generate renewable energy from land-intensive solar and wind projects and new transmission lines across states and tribal lands. Creating well-paid jobs for Californians without fancy college degrees in alignment with the national priority of re-shoring manufacturing of critical technologies and supplies to respond to global supply chain and national defense uncertainties is another priority doomed to CEQA pre-litigation and post-litigation bickering, costly multi-year studies, and uncertain judicial outcomes decided in decades, not months or years.

Legislative amendments to CEQA face significant political hurdles, and even if those hurdles are overcome, the amendments will not be effective unless judicial enforcement of CEQA is reshaped into traditional administrative law jurisprudence based on the Rule of Law.

A core principle of the United States, and other democratic governments globally, is that all people and all institutions are required to comply with the Rule of Law. As defined in Oxford Languages, the Rule of Law is “the restriction of the arbitrary exercise of power by subordinating it to well-defined and established laws.” As amplified by the World Justice Project, the rule of law requires the law to be “clear, publicized, and stable and to be applied evenly.” It also “ensures human rights as well as property, contract, and procedural rights.”

Courts can restore administrative law jurisprudence to CEQA jurisprudence by embracing the 1993 statutes and ignoring any substance or process not expressly required by the CEQA statutes of Guidelines. The Governor can revise the Guidelines to align with today’s civil rights, housing, environmental and economic justice, and climate priorities. The Legislature, and all the CEQA status quo defenders who lobby in the Legislature, need to recognize that the harms inflicted on California by weaponizing CEQA can far more effectively, equitably, and economically be achieved by statutes resolving policy disputes directly—not via CEQA.


431 Id.
Californians created our CEQA (and housing) mess, especially the Californians from my generation (Baby Boomers). We need to own this failure while also owning our environmental successes, like stripping tailpipe emissions of ninety-nine percent of smog pollution as confirmed by the EPA.\textsuperscript{432} We can do this, but not by talking past each other or refusing to talk with each other at all, or even by continuing to pretend that what we are doing with CEQA is to “protect the environment” instead of “protect my environment.”

The Foreign Tax Credit Redux

Bret Wells*

In the preamble to its 2022 final regulations, the Treasury Department provided multiple justifications for its amendments based on the historic policy goals of Section 901 and the manner that judicial case law has construed this provision. Yet, in fact, the amendments made by the 2022 final regulations deviate away from the historic policy goals of the U.S. foreign tax credit without any Congressional authorization for doing so. Moreover, these 2022 final regulations represent a strong repudiation of the Supreme Court’s own articulation of the Biddle doctrine in the PPL decision by attempting to formulate an interpretation of the Biddle doctrine that is inconsistent with the Supreme Court’s own interpretation of its own doctrine. The U.S. Treasury Department has forged a diametrically opposite policy approach in this era compared to the one that Congress chose to pursue in the circa 1918-1921 era when it enacted Section 901’s predecessor. In 1918, Congress adopted a unilateral foreign tax credit before a consensus on international taxation norms was forged, and the United States worked for a consensus on international norms in the succeeding years. In contrast, in 2022, the Treasury Department sought to deny foreign tax credit relief on destination-based taxes until a further international consensus on taxation of the digital economy is fully implemented. In 1918, Congress prioritized mitigation of international double taxation above the interests of the U.S. fisc and then worked to create a consensus on international taxation. In contrast, in 2022, the Treasury Department reversed the prioritization and created the real possibility of international double income taxation, which is antithetical to the policy goal that undergirds Section 901. Seen in light of its historical objectives and historical context, the Treasury Department’s amendments to its Section 901 final regulations fail to satisfy the text, purpose, and policy goals that guided the original enactment of the foreign tax credit regime.

* Bret Wells is the Law Foundation Professor of Law at the University of Houston Law Center. The author wishes to thank Haley Ritter, Rachel McMain, Sarah McMillin, and the rest of the Chapman Law Review editing team. The author also wishes to thank Christopher Dykes, reference and research librarian at the O’Quinn Law Library of the University of Houston Law Center, for his help in procuring titles. All views and potential errors in this article are solely the responsibility of the author.
Instead of pursuing the path that it has taken, the Treasury Department should withdraw its 2022 amendments.

The Treasury Department also did not address the policy implications of the implementation of the OECD Pillar Two framework even though the Treasury Department has endorsed that initiative. Under Pillar Two, a top-up tax would be applied to ensure that a minimum tax is paid by multinational enterprises, regardless of where they are headquartered or operate. These top-up taxes are conceptually an additional tax needed to arrive at a minimum tax and thus do not present a double taxation concern. As a result, these top-up taxes, applied by jurisdictions that adopt the OECD Pillar Two regime, should be denied foreign tax credit relief. In 2022, Congress adopted a corporate alternative minimum tax that does not comply with the GloBE rules. By enacting a provision that does not fit neatly with the GloBE rules, the enacted corporate minimum tax may represent a better outcome than if the United States had enacted a qualified IIR in compliance with the GloBE rules. But, the enacted legislation contains a deficiency. What should have been done concurrently with the enactment of this corporate alternative minimum tax (but was not done) was a companion amendment to Section 901 so that Section 901 would not afford foreign tax credit relief for any non-covered tax that is imposed as a top-up tax under the GloBE rules. Failing to do so has put the residual U.S. tax jurisdiction at risk of being eroded through minimum taxes imposed by other nations in preference to the corporate alternative minimum tax imposed by the United States. The OECD framework envisions that top-up taxes modelled after the GloBE rules would not be afforded foreign tax credit relief among nations, and so a denial of foreign tax credit relief for such top-up taxes by the United States would have been consistent with the international consensus endorsed by the OECD framework. Thus, the U.S. failure to make this conforming amendment to Section 901 represents a self-inflicted wound. Congress should correct this mistake by amending Section 901 to make it clear that top-up taxes under a qualifying IIR or a qualifying UTPR would not be afforded U.S. foreign tax credit relief. This is the reform that is needed under Section 901, not the imposition of close conformity requirements or the jurisdictional nexus requirements envisioned by the 2022 final regulations. Reform along these lines effectuates the policy goals sought by the OECD framework and also protects the U.S. tax base. It is now time for Congress and the Treasury Department to correct these mistakes.
INTRODUCTION

This has been a remarkable two years. In October 2021, the Organisation for Economic Co-operation and Development (“OECD”) and G-20 set forth a joint statement indicating their broad agreement on the Pillar One and Pillar Two recommendations of the OECD.1 As of November 4, 2021, 135 nations had agreed to make changes to their domestic tax laws to conform to the OECD Inclusive Framework.2 The OECD has since

---


issued model rules, commentary, and examples that set forth the design parameters for its Pillar Two proposal. And, draft model rules have been issued with respect to the nexus and revenue sourcing aspects of its so-called Pillar One proposal. Under the Pillar One draft, source jurisdictions would be afforded the right to assert taxation over remote sellers based on a formulaic reallocation of residual profits back to the market jurisdiction where the customer is located, even if the multinational enterprise lacked a permanent establishment in that country. Thus, the historical understanding of jurisdictional nexus and sourcing of gains would be modified if Pillar One’s proposal were implemented. The Biden administration has enthusiastically endorsed the OECD Inclusive Framework and has argued that its implementation is critical to all nations. The United Nations (“U.N.”) Committee of Experts on International Tax Matters initially provided comments to the Pillar One proposal set forth in the OECD Inclusive Framework. Nevertheless, this Committee


7 See id. at 2, 27.


9 See U.N. Comm. of Experts on Int’l Coop. in Tax Matters, Co-coordinators’ Paper on Tax Issues Related to the Digitalization of the Economy for the Twentieth Session of the
determined that it would set forth its own recommendation. It then set forth a new Article 12B to the U.N. model treaty that would allow countries to assert taxation over income generated from digital sales into a market jurisdiction. The U.N. approach in Article 12B allows market jurisdictions to tax income from automated digital services conducted with respect to its jurisdiction regardless of any physical nexus, and Article 12B allows the market jurisdiction to impose its tax on a gross basis, not a net basis. The U.N. explained that its new Article 12B seeks to preserve the domestic law taxing rights for States from which payments for automated digital services are made. The U.N. Model Treaty commentary notes that a significant minority of its members preferred the multilateral approach of the OECD Inclusive Framework in lieu of the bilateral approach that the U.N. Model Treaty affords. In 2022, perhaps in response to these concerns, the U.N. Committee of Experts on Tax Matters announced that it would commence work on a multilateral instrument that would allow nations to adopt its Article 12B proposal on a multilateral basis.


12 The commentary to the U.N. Model Convention explains that the constraints faced by developing countries’ tax administrations justifies allowing tax on a gross basis as that is an established means of collecting tax on nonresident persons. See id. at 24, para. 2; see also id. at 436, para. 5 (commentary on Article 12B). To address concerns about the imposition of a gross revenue tax on unprofitable companies, the commentary asserts that a modest gross rate of taxation of three or four percent would likely obviate this concern; but then it provides that taxpayers subject to the new article with the ability to opt for taxation on net profit at the rate applicable under domestic law. Taxpayers that exercise this option would be taxed on their “qualifying profits,” which the draft article defines as thirty percent of the beneficial owner’s consolidated automated digital business segment profitability ratio, multiplied by its gross automated digital services revenue in the jurisdiction. See id. at 24–25, para. 3; see also id. at 435, para. 4, 449–54, paras. 39–51 (commentary on Article 12B).

13 See id. at 24, para. 2.

14 See id. at 436–40, paras. 8–16 (commentary on Article 12B).

Yet, amid these global developments, the U.S. Treasury Department issued final regulations in 2022 under Section 901, providing, for the first time, that a foreign levy is eligible for foreign tax credit relief only if the foreign taxing jurisdiction (i) utilizes jurisdictional nexus and sourcing rules that closely conform to the existing U.S. rules (referred to as an “attribution requirement”)\(^\text{16}\) and (ii) calculates its tax base in a manner that closely conforms to the current U.S. statutory provisions.\(^\text{17}\) The 2022 final regulations represent a substantial reformulation of the U.S. foreign tax credit eligibility standards. Later in 2022, Congress enacted a new fifteen percent corporate alternative minimum tax as part of the Inflation Reduction Act of 2022.\(^\text{18}\)

Unfortunately, for reasons that will be further addressed in Part I, the Treasury Department’s regulatory amendments represent an inappropriate departure from the text, purpose, and policy goals of the historic mission of the U.S. foreign tax credit. A careful review of the 2022 final regulations in light of the legislative history, case law, and statutory text of Section 901 is addressed in Part I.

In Part II, this article will address the changes to Section 901 that should have been made at the time that Congress enacted a new corporate alternative minimum tax as part of the Inflation Reduction Act of 2022. The OECD Pillar Two proposal envisions jurisdictions will impose top-up taxes in accordance with the GloBE rules set forth in the OECD Pillar Two project. The imposition of these top-up taxes poses a normative design challenge to the U.S. foreign tax credit regime because these top-up taxes should not reduce the U.S. taxation on foreign income, but should instead be imposed in addition to the U.S. tax. The OECD framework agrees that these top-up taxes should not be afforded foreign tax credit relief so as to reduce any covered taxes like the U.S. taxation of Controlled Foreign Companies (“CFCs”) or its own domestic minimum tax on U.S. income. Nevertheless, current U.S. law does not envision the concept of a “top-up” tax, and thus, at present, such taxes are afforded U.S. foreign tax credit relief under existing law. It is here where reform of Section 901 should have been done but was not. For the reasons discussed in Part II, Congress should amend Section 901 to deny foreign tax credit relief for any top-up tax enacted in accordance with the GloBE rules.

\(^{16}\) See Treas. Reg. § 1.901-2(b)(5) (as amended in 2022).
\(^{17}\) See id. §1.901-2(b)(4).
\(^{18}\) See I.R.C. § 55(b)(2) (West 2022).
I. CONSIDERATION OF THE TREASURY DEPARTMENT’S AUTHORITY FOR THE 2022 FINAL REGULATIONS

What motivated the regulatory shift set forth in the 2022 final Treasury Regulations? In 2020, the Treasury Department announced that it was concerned that several foreign countries had adopted, or were considering adopting, a variety of novel extraterritorial taxes (such as digital services taxes, diverted profits taxes, or equalization levies).19 The Treasury Department then indicated that these taxes diverge significantly from traditional norms of international taxation and thus raise considerable policy concerns as to whether they represent “an income tax in the U.S. sense.”20 The Treasury Department repeated this concern in the preamble to its 2022 final regulations.21

Sandwiched between these 2020 and 2022 regulatory statements is a Treasury Department statement from October 2021 where it announced that the United States and its major trading partners had agreed to new international taxation norms for remote sellers in the OECD’s Pillar One proposal.22 Within one month of the issuance of the 2022 final regulations, the OECD issued a discussion draft for how these new nexus and sourcing rules would reattribute a portion of remote sellers’ profits to the country where their customers were located.23 Thus, on the one hand, the Treasury Department, in the preamble to its 2022 final regulations, asserts that taxation based on customer location violates existing “international norms,” but on the other hand, the Treasury Department has signed-on to an OECD initiative that reformulates the nexus and sourcing rules to do exactly that.24

What is going on? Why has the Treasury Department agreed to extraterritorial taxation under the rubric of the OECD Inclusive

24 This observation has been pointedly made by a large group of U.S. multinational companies. See Letter from All. for Competitive Tax’n to Sec’y Janet Yellen, Dep’t of the Treasury (Feb. 24, 2022), http://www.actontaxreform.com/media/gpuh55nj/act-letter-to-treasury-2021-final-ftc-reges_20220224-final.pdf [http://perma.cc/9YF5-8HHB].
Framework on the one hand, but then issued 2022 final regulations that, on the other hand, deny foreign tax credit relief for such assertions of extraterritorial taxation over remote sellers by market jurisdictions? It has been speculated that the Treasury Department is using its regulatory authority to pressure other countries into joining the OECD framework as a precondition for the United States to provide foreign tax relief.25 This view is given some credence due to the fact that Treasury Department officials and the Treasury Department’s preamble explanation both have stated that implementation of the Pillar One framework would likely require an immediate amendment to its newly issued 2022 final regulations.26 If this is the explanation, then this raises the prospect that the 2022 final regulations are being used as a bargaining chip to promote policy goals other than those of the statutory provision to which they were promulgated under. Congress has endorsed neither the OECD initiative27 nor the Treasury Department’s usage of the foreign tax credit eligibility standards of Section 901 as a bargaining chip in that broader multilateral negotiation.28 As a result, significant commentary has already questioned whether these regulations would pass muster under ‘Chevron’s deference standards,29 and, in fact, it appears

26 See, e.g., Stephanie Soong Johnston, U.S. to Mull Credit for Qualified Domestic Minimum Taxes, 105 TAX NOTES INT’L 1572 (Mar. 24, 2022) http://www.taxnotes.com/tax-notes-today-federal/credits/us-mull-credit-qualified-domestic-minimum-taxes/2022/03/24/7d9r8. Moreover, the preambles to both the proposed and final regulations make clear that countries that adopt virtual taxation under the OECD Inclusive Framework may necessitate amendment of the jurisdictional nexus standards. See T.D. 9959, 87 Fed. Reg. 276, 288 (Jan. 4, 2022); see also Notice of Proposed Rulemaking, REG-101657-20, 85 Fed. Reg. 72078, 72089 (Nov. 12, 2020) (“If an agreement [on the OECD Inclusive Framework] is reached that includes the United States, the Treasury Department recognizes that changes to the foreign tax credit system may be required at that time.”).
27 In fact, at present, there is significant concern among some in Congress that the OECD framework, as currently formulated, does not adequately address the U.S. fiscal interests and that it would place the United States at a competitive disadvantage. See Letter from Assemb. Kevin Hern et al., to Sec’y Janet Yellen, Dep’t of the Treasury (Jan. 19, 2022), http://hern.house.gov/uploadedfiles/hern-oecd-letter.pdf [http://perma.cc/WD7P-45D4]; see also Letter from U.S. Senate Comm. on Fin. to Sec’y Janet Yellen, Dep’t of the Treasury (Feb. 16, 2022), http://www.finance.senate.gov/imo/media/doc/finance_republicans_oecd_follow-up.pdf [http://perma.cc/R9CL-E2BM].
29 See Chevron U.S.A. Inc. v. Nat. Res. Def. Council, Inc., 467 U.S. 837 (1984) (stating that “if the intent of Congress is clear, that is the end of the matter, for the court, as well as the agency, must give effect to the unambiguous expressed intent of Congress” [so-called Chevron Step 1]; but, if Congress has left a gap to fill then the Court looks to whether the regulation is a permissible construction of the statute [so-called Chevron Step Two]; see also Gary B. Wilcox & Lucas Giardelli, Will Jurisdictional Nexus Survive Chevron Step 1? 174 TAX NOTES FED. 1379, 1380 (2022).
that at least one firm is proceeding to make such a challenge. In response to these deference concerns, the Treasury Department has strongly denied that the OECD initiative had any bearing on its decision to issue its 2022 final regulations, and the Treasury Department has reaffirmed its belief that its regulatory revisions represent a faithful interpretation of Section 901’s eligibility requirements. In this part of the article, a singular question is addressed: do the amendments made in the 2022 final regulations represent a faithful interpretation of Section 901’s eligibility requirements? After addressing this question, this article then addresses the appropriate reform that remains to be done with respect to the U.S. foreign tax credit in Part II.

A. Text, Purpose, and Policy Goals of Section 901

In the preamble to its 2022 final regulations, the Treasury Department provided multiple justifications for its issuance of its 2022 final regulations. One asserted justification is that the statutory text, purpose, and policy goals of Section 901 support the issuance of its new 2022 final regulations, as demonstrated by the following excerpt:

[T]he Treasury Department and the IRS’s determination that regulations are necessary and appropriate to ensure that the U.S. fisc does not bear the costs of such taxes derives from the text, purpose, and policy of section 901, and not from any foreign policy goals. The Treasury Department and the IRS have concluded that these novel extraterritorial taxes (some of which are currently in force and being levied on U.S. taxpayers) are contrary to the text and purpose of section 901 and therefore must be addressed now.

The above statement makes clear that it is the Treasury Department’s position that its 2022 final regulations are a faithful interpretation of the law derived from the text, purpose, and policy

---


31 The preamble to its final regulations includes this rebuttal: The Treasury Department and the IRS agree that international forums can be an effective way of discouraging foreign jurisdictions from enacting extraterritorial taxes; indeed, the Treasury Department is actively engaged in and supporting negotiations under the auspices of the Inclusive Framework that would result in their elimination. However, contrary to the comments’ assertion, the Treasury Department and the IRS’s determination that regulations are necessary and appropriate to ensure that the U.S. fisc does not bear the costs of such taxes derives from the text, purpose, and policy of section 901, and not from any foreign policy goals.


32 Id.

33 Id.
goals of Section 901, but the statement goes further by asserting a clear priority rule: if there is a risk of double taxation arising from tax levies that are inconsistent with jurisdictional nexus and sourcing conformity norms of the United States, then the protection of the United States fisc takes precedence over the goal of mitigating against double international taxation. Is this prioritization consistent with the prioritization that Congress intended when it enacted Section 901? To answer that question, the text, purpose, and policy goals of Section 901 must be examined. This article addresses that inquiry in this part.

The Treasury Department also offered the following additional rationale for its new jurisdictional nexus and sourcing conformity requirement:

The foreign tax credit is not intended to subsidize foreign jurisdictions at the expense of the U.S. fisc.

... [T]he fundamental purpose of the foreign tax credit—to mitigate double taxation with respect to taxes imposed on income—is served most appropriately if there is substantial conformity in the principles used to calculate the base of the foreign tax and the base of the U.S. income tax. This conformity extends not just to ascertaining whether the foreign tax base approximates U.S. taxable income determined on the basis of realized gross receipts reduced by allocable costs and expenses, but also to whether there is a sufficient nexus between the income that is subject to tax and the foreign jurisdiction imposing the tax. Therefore, the final regulations retain the requirement in the 2020 FTC proposed regulations that for a foreign tax to qualify as an income tax, the tax must conform with established international jurisdictional norms, reflected in the Internal Revenue Code and related guidance, for allocating profit between associated enterprises, for allocating business profits of nonresidents to a taxable presence in the foreign country, and for taxing cross-border income based on source or the situs of property.34

This jurisdictional nexus and sourcing conformity requirement will be addressed concurrently with the analysis of the text, purpose, and policy goals of Section 901 in this part.

The Treasury Department, in the alternative, also argued that its regulatory changes are supported by the evolving case law that has interpreted Section 901, as the following statement so indicates:

Judicial decisions and administrative guidance over the past century have interpreted the term “income, war profits, and excess profits tax,” which is not defined in section 901 or by the limited initial explanation in the early legislative history. These interpretations have consistently followed the principle, introduced by the Biddle court, that the determination of whether a foreign tax is creditable under section 901 is

34 Id. at 284–85.
made by evaluating whether such tax, if enacted in the United States, would be an income tax (in other words, whether the foreign tax is “an income tax in the U.S. sense”). See *PPL Corp. v. Comm’r*, 569 U.S. 329, 335 (2013). See also *Inland Steel Co. v. United States*, 230 Ct. Cl. 314, 325 (1982) (“Whether a foreign tax is an income tax under I.R.C. 901(b)(1) is to be decided under criteria established by United States revenue laws and court decisions.”).

... It is appropriate for the definition of a creditable tax to incorporate the concept of jurisdictional nexus from the U.S. tax law. The fact that U.S. tax rules have changed since the foreign tax credit provisions were first enacted does not preclude an interpretation of the term “income tax” to reflect U.S. norms, because the principle of “an income tax in the U.S. sense” incorporates an evolving standard of what constitutes an income tax in the U.S. sense.35

So, are the changes made by the 2022 final regulations a faithful distillation of the holdings in the case law as interpreted in *Biddle*, *PPL*, and *Inland Steel* as the Treasury Department has asserted, or do these regulatory changes repudiate those holdings?36 The principles utilized in the case law and how those principles align with the principles set forth in the regulations are addressed in Part I.B. This article then addresses the U.S. treaty implications arising from the 2022 final regulations in Part I.C.

As introduced in the Introduction, the 2022 final regulations set forth a jurisdictional nexus and sourcing conformity requirement under the rubric of a so-called attribution requirement.37 Under the attribution requirement, the foreign tax base, as applied to a nonresident person, must be limited to the activities conducted from within the foreign country (including functions, assets, and risks located in foreign country) without taking into account destination-based criteria (e.g., location of customers, users, or persons from whom a nonresident person makes purchases in the foreign country).38 In addition, the foreign tax base, as applied to a nonresident person, must be limited in its taxation of gains to only those gains that arise from the sale or disposition of real property located in the foreign country (or interest in resident corporation or other entity that owns such property) under rules similar to those of the Foreign Investment in Real Property Tax Act (“FIRPTA”). Alternatively, the tax base must be limited to the taxation of business property that is part of

---

35 *Id.* at 283.
37 See Treas. Reg. § 1.901-2(b)(5) (as amended in 2022); *see also* Treas. Reg. § 1.903-1(c)(2)(iii) (as amended in 2022).
taxable presence in the foreign country (including interests in partnership or pass through entity attributable to such property) as determined under principles similar to those of Section 864(c). In addition, in order for a foreign tax to qualify as an “in lieu of” tax under Section 903, the foreign jurisdiction must have an income tax that utilizes jurisdictional nexus and sourcing rules that closely conform to existing U.S. law, and the imposition of the “in lieu of” tax must separately utilize sourcing rules that closely conform to the sourcing rules found in existing U.S. tax laws. As a result of these new jurisdictional and sourcing conformity requirements, a withholding tax on services or royalties based on the residence of the payor would fail to be creditable as that sourcing rule does not conform to the U.S. sourcing rule for services or for royalties.

This is so whether or not services were actually performed outside that foreign jurisdiction. Furthermore, because several countries attempt to impose taxation on capital gains of nonresident persons other than gains attributable to real property, their foreign tax levies would not be eligible for foreign tax credit relief under the 2022 final regulations.

41 See Treas. Reg. § 1.901-2(b)(5)(i)(B)(1)–(2) (as amended in 2022). The sourcing for income derived from services is the place of service whereas income from royalties is generally sourced by the place of use of the intangible asset. See also Treas. Reg. § 1.903-1(c)(2)(iii) (as amended in 2022). The Treasury Department, in proposed regulations, has provided a limited exception to this sourcing conformity requirement for withholding taxes on royalties when those withholding taxes are paid pursuant to a single-country license agreement. See Prop. Treas. Reg. § 1.903-1(c)(2). Under this single-country license agreement exception, a withholding tax on gross royalties imposed on a nonresident remains a covered tax eligible for U.S. foreign tax credit relief if the income is earned pursuant to a single country license agreement. See id. The proposed regulations contemplate two varieties of a single-country license agreement that can satisfy this requirement. Under the first variation, a royalty payment made under a single country license agreement could qualify if each of the following requirements are satisfied: (i) the royalty payment is made pursuant to a license agreement, (ii) the royalty payment is characterized as a royalty in the foreign tax jurisdiction, and (iii) the agreement limits the territory of use for the intangible property to the country that imposes the withholding tax for the royalty subject to taxation in that jurisdiction. See id. Under a second variation, a multi-jurisdictional license agreement could satisfy the single-country license exception if requirements (i) and (ii) in the immediately preceding sentence were satisfied and the agreement does not misstate the territory of use for the intangible property and the amount of the royalty under the license agreement complies with the arm’s length standard. See id. § 1.903-1(c)(2)(iv)(B). For a further analysis of this limited concession to one category of income (namely, royalty income), see Joseph Isenbergh & Bret Wells, International Taxation: U.S. Taxation of Foreign Persons and Foreign Income ¶¶ 56.11.4 (Wolters Kluwer 6th ed. 2024) (forthcoming) (on file with author).
42 It should be noted that this aspect of the regulations has created considerable public comment. The Treasury Department has indicated that it may ameliorate this outcome in future guidance, but as of this article’s submission, such guidance had not been provided.
In this part, the text, purpose, and policy goals of Section 901 are considered in light of this newfound jurisdictional nexus and sourcing conformity requirement. As previously mentioned in the Introduction, the government in its preamble asserted that the text, purpose, and policy goals of Section 901 support the Treasury Department’s imposition of a jurisdictional nexus and sourcing conformity requirement into its 2022 final regulations. However, it is important to understand that the preamble to the 2022 final regulations contains a contradictory statement on this score. In this regard, the Treasury Department made the following admission about the text, purpose, and policy goals of Section 901 elsewhere in the very same preamble explanation:

Congress has not explicitly addressed jurisdictional nexus with respect to the foreign tax credit. There is no statutory provision that addresses whether the foreign tax credit should be allowed for taxes imposed outside of traditional U.S. taxing norms. . . . The statute is silent with respect to jurisdictional nexus, and it is reasonable and appropriate for regulations to apply U.S. tax concepts in addressing the creditability of extraterritorial foreign levies that Congress could not have anticipated when the foreign tax credit provisions were first enacted.43

A court will need to determine what weight, if any, should be placed on the Treasury Department’s reliance on the text, purpose, and policy goals of Section 901 as the asserted justification for its regulatory changes when the Treasury Department elsewhere makes the contradictory statement that the statutory text and the legislative purpose are silent with respect to these elements of its 2022 final regulations.

But even so, this still leaves unanswered the question of whether or not the actual statutory text and legislative history are in fact silent on this question. The Treasury Department indeed is correct in its observation that Section 901 in its current form has no explicit jurisdictional nexus or sourcing conformity requirement. So, is this omission because Congress did not want to impose one, or is this omission explained (as the Treasury Department has alleged) by the fact that “Congress could not have anticipated” this matter? To resolve that question, one must look to the facts and circumstances that existed at the time of Section 901’s enactment.

To begin, it is important to note that the U.S. tax laws initially did not provide for foreign tax credit relief under the tax laws of 1909 and 1913,44 but the income tax rates were substantially lower in that era, so the cost of not providing foreign tax credit relief was

---

However, with the advent of World War I, tax rates increased sharply in the United States and other countries. With increasing tax rates in both foreign countries and the United States, the cost of international double taxation became significant to U.S. multinationals and represented a threat to international trade. International double taxation would be the result if both the host country and the United States asserted simultaneous taxing jurisdiction over the same foreign income.

When Section 901 was enacted in 1918 (effective starting in 1919), it read as follows: “the amount of any income, war-profits and excess-profits taxes paid during the taxable year to any foreign country, upon income derived from sources therein, or to any possession of the United States . . . .” Prior to 1921, the United States had not articulated a comprehensive concept of source, but the above italicized language suggests that a jurisdictional nexus and sourcing conformity requirement initially was part of the foreign tax credit eligibility requirements, but not for long. In 1921, Congress enacted the foreign tax credit limitation regime as the means to address the over-crediting of foreign tax credits so that the foreign tax credit could not be used against U.S. source income but could be used against any foreign source income. Simultaneously with the

45 See Stanley S. Surrey, The United States Taxation of Foreign Income, 1 J.L. & ECON. 72, 73 n.3 (1958) (noting that in 1917, war time income tax rate increases were adopted).
46 See id. Thomas Sewell Adams, viewed as the architect of the foreign tax credit, stated: “[i]n the midst of the war, when the financial burden upon the United States was greater than it had ever been, I proposed to the Congress that we should recognize the equities . . . by including in the federal income tax the so-called credit for foreign taxes paid . . . .” Thomas S. Adams, International and Interstate Aspects of Double Taxation, 22 NAT’L TAX ASS’N PROC. 193, 198 (1929).
47 See Surrey, supra note 46, at 73; see also H.R. Rep. No. 65-767 (1918), 1939-2 C.B. 86 (explaining the rationale and legislative history for a foreign tax credit); see also id. at 93 (“With the corresponding high rates imposed by certain foreign countries that taxes levied in such countries in addition to the taxes levied in the United States upon citizens of the United States place a very severe burden upon such citizens.”).
50 See id.
51 See Revenue Act of 1921, Pub. L. No. 67-98, §§ 222(a)(5), 238(a), 42 Stat. 227, 249, 258. Although not further discussed in this article, this limitation regime has taken various forms. In 1932, Congress decreed that taxpayers were required to use the lesser of an overall or per-country limitation. See Revenue Act of 1932, Pub. L. No. 72-154, ch. 209, § 131(b), 47 Stat. 169, 211. In 1954, the overall limitation was repealed and only the per-country limitation regime existed. See I.R.C. § 904 (West). In 1960, taxpayers were given the option to use either a per-country or an overall limitation computation. See Act of Sept. 14, 1960, Pub. L. No. 86-780, § 1(a), 74 Stat. 1010. In 1976, the per-country limitation was
The enactment of Section 904’s predecessor, Congress removed any jurisdictional nexus or sourcing conformity requirement from Section 901’s predecessor so that it would read as follows: “the amount of any income, war-profits and excess-profits taxes paid during the same taxable year to any foreign country, or to any possession of the United States.”

Finally, Congress’ decision to affirmatively remove any reference to sourcing or jurisdictional nexus from Section 901(b)(1)’s predecessor occurred at the very same time that Congress, elsewhere in the 1921 Tax Act, expanded the complexity and specificity of the United States’ own sourcing rules and directed the Commissioner of Internal Revenue to develop apportionment rules for U.S. expenses. In 1932, Congress made it explicitly clear that the rules for allocating and apportioning U.S. expenses between U.S. source income and foreign source income must be utilized in order to determine the foreign tax credit limitation so that it applied on a foreign source net income basis.

Thus, Congress thought a lot about the interaction of sourcing rules, the foreign tax credit limitation, and the scope of the foreign tax credit eligibility criteria during this era. Additionally, Congress affirmatively walked back any indication that it would require a jurisdictional nexus and sourcing conformity requirement for foreign tax credit eligibility purposes even though the original statutory provision had such a requirement. Instead of continuing to impose a jurisdictional nexus and sourcing conformity rule, Congress confined its income as if reading naturally.


54 See Revenue Act of 1932, Pub. L. No. 72-154, § 131(e), 47 Stat. 169, 212 (adding an explicit cross-reference to the sourcing rules of former Section 119, which was the predecessor to Section 861). Courts have held that the sourcing rules are utilized to determine the foreign tax credit limitation. See, e.g., Int’l Standard Elec. Corp. v. Comm’r, 1 T.C. 1153 (1943), aff’d, 144 F.2d 487 (2d Cir. 1944).
sourcing and expense apportionment rules to the foreign tax credit limitation computation.

The removal of any jurisdictional nexus and sourcing rule, along with the concurrent adoption of the predecessor to Section 904 and the concurrent enactment of detailed sourcing rules all at the same time, demonstrates that Congress recognized that its sourcing rules were important but limited their application to the foreign tax credit limitation context. The current version of Section 901 contains the added words “or accrued,” but otherwise Section 901 has remained unchanged in relevant part since 1921 as can be seen in the following redline version: “the amount of any income, war profits, and excess profits taxes paid or accrued during the same taxable year to any foreign country, upon income derived from sources therein, or to any possession of the United States.”

Thus, the evolution of the statutory text makes it clear that Congress originally included a jurisdictional sourcing conformity requirement, but Congress soon thereafter eliminated that requirement as an affirmative change in the statutory text. Prior to 2022, the Treasury Department accepted the premise that no sourcing conformity requirement existed under Section 901 as the Treasury Department’s prior 1983 final regulations explicitly granted foreign tax credit relief to a foreign levy that would have failed a sourcing conformity requirement. Thus, at least two conclusions from the statutory text become clear. First, the text and legislative history demonstrate that Congress thought about sourcing in the context of Section 901, and when it did so it removed it from Section 901 as an affirmative action. And second, the Treasury Department’s prior regulations explicitly recognized that Section 901 had no sourcing conformity requirement through the end of 2021.

The Treasury Department’s preamble to its 2022 amendments attributes the statutory silence on jurisdictional nexus to the fact that Congress could not have anticipated objectionable assertions of jurisdictional nexus in the circa 1918-

55 See I.R.C. § 901(b) (West 2022).
56 A hat tip is owed to Robert Culbertson who first provided this redline version of Section 901 to demonstrate how it was amended in relevant part since its inception. See Culbertson, supra note 28, at 186 n.6 (emphasis in original).
57 See Treas. Reg. 1.903-1(b)(3) (example 3), 48 Fed. Reg. 46272, 46296 (Oct. 12, 1983). In the regulatory example, a withholding tax was applied on technical services performed outside the jurisdiction that applied the withholding tax. Nevertheless, the 1983 regulations concluded that the withholding taxes were creditable. For a further discussion of this example, see Wilcox & Giardelli, supra note 29, at 1157.
The Foreign Tax Credit Redux

1921 era.\textsuperscript{58} As previously discussed, this assertion is contradicted by the actual legislative history attendant with the original enactment and subsequent amendment of Section 901’s predecessor in 1918 and 1921, respectively, and by Congress’ attention to source rules in that era. In addition, this assertion is contradicted by the jurisdictional practices that existed at the time of Congress’ enactment of Section 901’s predecessor. In this regard, the foreign tax credit was enacted into U.S. law during a turbulent period when international taxation norms were not agreed on, to say the least. In the post-World War I era, nations faced crushing war debt, and, at that time, there was a substantial increase in income taxation measures around the world that threatened international trade. Formulary apportionment among nations, without the prerequisite of a permanent establishment, was the treaty norm at that time.\textsuperscript{59} There was no shared understanding of jurisdictional norms, nor was there any shared understanding of commonly accepted transfer pricing methods.\textsuperscript{60}

Prior to the work of the League of Nations, there was a divergence in how nations asserted taxation over profits arising from economic activities that had some connection to more than one jurisdiction. It was for this reason that the League of Nations commenced its work in 1923 to develop standards for how international taxation should occur.\textsuperscript{61} A consensus on jurisdictional norms and on transfer pricing practices was not

\textsuperscript{58} Specifically, the Treasury Department made this assertion in its preamble to the final regulations: “[t]he statute is silent with respect to jurisdictional nexus, and it is reasonable and appropriate for regulations to apply U.S. tax concepts in addressing the creditability of extraterritorial foreign levies that Congress could not have anticipated when the foreign tax credit provisions were first enacted.” T.D. 9959, 87 Fed. Reg. 276, 284 (Jan. 4, 2022).

\textsuperscript{59} See Mitchell B. Carroll, A Brief Survey of Methods of Allocating Taxable Income Throughout the World, in LECTURES ON TAXATION 131, 151–53, 168–70 (Roswell Magill ed., 1932) (stating that fractional apportionment was the primary method of resolving double taxation for Spain and Switzerland and was also used by France; also providing an analysis of how Austria, Czechoslovakia, Hungary, and Poland had all formulated significant apportionment methodologies); see also John G. Herndon, RELIEF FROM INTERNATIONAL INCOME TAXATION: THE DEVELOPMENT OF INTERNATIONAL RECIPROCITY FOR THE PREVENTION OF DOUBLE INCOME TAXATION 15 (1932) (describing a pre-existing German-Holland treaty where income apportionment was used for a railroad between the two countries); see also Edwin A. Seligman, DOUBLE TAXATION AND INTERNATIONAL FISCAL COOPERATION 138 (1928) (recognizing that Great Britain had employed formulary apportionment methods with respect to its colonies).


\textsuperscript{61} This author has elsewhere discussed in depth the development of international norms during this era which can be reviewed for further detail on topics that support this understanding of the historical evolution of international taxation among nations. See Bret Wells & Cym Lowell, Tax Base Erosion and Homeless Income: Collection at Source is the Linchpin, 65 Tax L. Rev. 555 (2012) [hereinafter Wells & Lowell, Linchpin]; Bret Wells & Cym Lowell, Income Tax Treaty Policy in the 21st Century: Residence vs. Source, 5 COLUM. J. TAX L. 1 (2013).
reached until the mid-1930s. The concept of limiting a jurisdiction’s taxation to business profits attributable to a permanent establishment was not an accepted norm until the League of Nations model treaty released in draft form in 1935. The League of Nations work transformed the world, but that transformation happened in the decades after the U.S. foreign tax credit had been enacted into U.S. law.

Rather, it was in the midst of the chaos, not after an international consensus had been formed, when Congress enacted the foreign tax credit. The foreign tax credit came into U.S. law as a measure to bolster international trade and protect U.S. persons from the threat of international double income taxation in an era when it was unclear whether or not an international consensus on jurisdictional taxation norms could be forged. The foreign tax credit is a unilateral measure adopted by the U.S. tax system. Although the credit represents a deference by the U.S. Treasury to the taxing power of foreign countries, it does not arise by treaty. On the surface at least, the U.S. tax system asks nothing from foreign treasuries in return for the credit. The case law has long recognized the Congressional goal of Section 901: the foreign tax credit should prevent worldwide double income taxation on the same foreign profits. The U.S.

---


63 See HERNDON, supra note 59, at 42.

64 This historical evolution is addressed in detail in Wells & Lowell, Linchpin, supra note 61.

65 T.S. Adams would later make these prescient observations:

In the midst of the war, when the financial burden upon the United States was greater than it had ever been, I proposed to the Congress that we should recognize the equities which I have just noted, by including in the federal income tax the so-called credit for foreign taxes paid . . . I had no notion, ladies and gentlemen, when I proposed it, that it would ever receive serious consideration. I expected it to be turned down with the reply which I have received so often from legislative committees: “Oh, yes, Doctor, that is pretty good, but the finances won’t permit it.” But to my surprise, the credit for foreign taxes was accepted and approved, because it touched the equitable chord or sense, and because double taxation under the heavy war rates might not only cause injustice but the actual bankruptcy of the taxpayer.

See Adams, supra note 46, at 198.

foreign tax credit provides relief for any income and excess profits\textsuperscript{67} taxes paid or accrued to a foreign country. Providing a credit against the U.S. tax liability ensures that U.S. taxation would not be applied to the extent that a foreign country already asserted taxation over that income, so that a single level of international taxation would apply. For example, in 1892, the Netherlands adopted a tax credit for traders deriving income from the then Dutch East Indies.\textsuperscript{68} “The United Kingdom in 1916 granted a partial tax credit to traders who had paid taxes to other territories of the [British] empire.”\textsuperscript{69} However, the United States apparently was the first country to adopt a broad-based foreign tax credit that would be granted to any nation’s income tax.\textsuperscript{70} The creation of a broad-based foreign tax credit was principally the invention of Thomas S. Adams, an economic advisor to the Treasury Department at the time.\textsuperscript{71} T.S. Adams explained the rationale for the enactment of Section 901 in the following terms:

\begin{quote}
221 F.2d 134, 137 (2d Cir. 1955) (“The primary objective of [the foreign tax credit regime] is to prevent double taxation and a secondary objective is to encourage American foreign trade.”). The legislative history is consistent and longstanding. See also H.R. REP. NO. 83-1337, at 76 (1954) (“The [foreign tax credit] provision was originally designed to produce uniformity of tax burden among United States taxpayers, irrespective of whether they were engaged in business in the United States or engaged in business abroad.”); S. REP. NO. 73-558, at 39 (1954) (“The present [foreign tax] credit . . . does relieve the taxpayer from a double tax upon his foreign income.”); H.R. REP. NO. 65-767, at 11 (1918) (explaining the rationale for a foreign tax credit, the legislative history stated as follows: “[w]ith the corresponding high rates imposed by certain foreign countries that taxes levied in such countries in addition to the taxes levied in the United States upon citizens of the United States place a very severe burden upon such citizens.”); Joint Comm. on Tax’n, 99th Cong., General Explanation of the Tax Reform Act of 1986 at 852 (Comm. Print 1986).

\textsuperscript{67} Excess profits taxes were imposed on only a portion of total income in excess of a given rate of return. See Income, Excess Profits, and Estate Taxes: Hearings Before the H. Comm. on Ways and Means, 65th Cong. 15 (1918) (statement of W.G. McAdoo, Treasury Secretary) (“By an excess-profits tax we mean a tax upon profits in excess of a given return on capital.”); see also George E. Holmes, Federal Income Tax, War-Profits and Excess-Profits Taxes 136 (1919) (stating excess profits taxes were imposed on only a portion of total income). The statute also refers to “war profits” taxes. See I.R.C. § 901 (West 2022). For a historical definition of a war profits tax, see Income, Excess Profits, and Estate Taxes: Hearings Before the H. Comm. on Ways and Means, 65th Cong. 15 (1918) (statement by W.G. McAdoo, Treasury Secretary) (“By a war-profits tax we mean a tax upon profits in excess of those realized before the war.”). By World War II, war-profits taxes were viewed as simply a subcategory of excess profits taxes. See Kenneth James Curran, Excess Profits Taxation 2 (American Council on Public Affairs 1943) (stating that “the term ‘excess profits tax’ [today is used] to describe any levy that is confined to a segment of a taxpayer’s income that is considered excessive, no matter by what standard of measurement it is determined”). Thus, for clarity to the modern reader, this paper discusses income taxes and excess profits taxes.

\textsuperscript{68} See Stanley S. Surrey, Current Issues in the Taxation of Corporate Foreign Investment, 56 COLUM. L. REV. 815, 818 n.4 (1956).

\textsuperscript{69} Id.

\textsuperscript{70} See id.

\textsuperscript{71} See Graetz & O’Hear, supra note 48, at 1038–39 n.71; Bret Wells, The Foreign Tax Credit War, 2016 BYU L. REV. 1895, 1900 n.11 [hereinafter Wells, Foreign Tax Credit War] (2016); Culbertson, supra note 28, at 170.
There is something in the legislative mind which recognizes that if one taxpayer is being taxed twice while the majority of men similarly situated are being taxed only once, by the same tax, something wrong or inequitable is being done which, other things being equal, the legislator should correct if he can.\textsuperscript{72}

Imposing a close conformity requirement would have impeded, not promoted, the purpose and policy goals of the foreign tax credit in the very era in which the U.S. enacted the foreign tax credit. The adoption of the foreign tax credit afforded deference to developing and developed nations alike, as it recognized the foreign jurisdiction’s primary right to assert taxation over foreign income first at a time when there was no consensus on international taxation norms. As T.S. Adams observed,

The explanation is simple. Every state insists upon taxing the non-resident alien who derives income from sources within that country, and rightly so, or at least inevitably so. . . . But [the average state] refuses to recognize when one of its own citizens or nationals gets income from a foreign source that he inevitably will be taxed abroad. As a necessary corollary of the principle of taxing at source or origin which it has adopted, the home state owes an exemption of some kind to its own citizen or resident who derives income from a foreign source or sources.\textsuperscript{73}

Prevention of double taxation, in short, calls for a self-denying ordinance in the home state — rather than concessions from a foreign state. The state of domicile must protect its own residents.\textsuperscript{74}

It is important to note where T.S. Adams indicated protection was needed. The protection that was needed was against double taxation of U.S. persons, not the revenue interests of the U.S. fisc. The revenue consequences to the U.S. fisc were subordinated to the more important priority of protecting against international double income taxation on foreign income earned by U.S. persons.\textsuperscript{75} The enactment of Section 901’s predecessor placed a higher priority on the goal of avoiding international double taxation and, as a collateral consequence, subordinated the financial interests of the U.S. fisc. Of course, T.S. Adams would work as the U.S. representative to the League of Nations to forge an international consensus on norms of international taxation, but there was no assurance that such a consensus would be obtained and in fact none was obtained during his lifetime.

\textsuperscript{72} Adams, \textit{supra} note 46, at 197.
\textsuperscript{73} Id. at 192–99.
\textsuperscript{74} T.S. Adams, \textit{Interstate and International Double Taxation, in Lectures on Taxation: Columbia University Symposium 101} (1932).
\textsuperscript{75} In the preamble to the 2022 amendments, the Treasury Department reverses the priority ordering rule such that it places a higher priority for protecting the interests of the U.S. fisc over the goal of preventing international double income taxation to U.S. persons. \textit{See} T.D. 9959, 87 Fed. Reg. 276, 284–85 (Jan. 4, 2022).
Thus, when push came to shove, the most important policy goal that Congress chose to prioritize as of first importance was protecting U.S. persons from international double income taxation. This is seen through Congress’ unilateral enactment of Section 901’s predecessor and the removal of a sourcing conformity requirement from the statutory text of that provision because the goal of protecting against international double income taxation took priority. The interests of the U.S. fisc took a back seat. Again, this was an amazing act of statesmanship on the part of the United States during an era of great international chaos. At the end of World War I, the United States believed that continued international trade among nations was critical to a future peaceful world and wanted to ensure that international double taxation did not inhibit the return of peaceful trade.76 A return to international comity was paramount. Instead of using the threat of double taxation as a bargaining chip with other nations in the post-war era, the United States gave a preference to the goal of preventing international double income taxation. Thus, a careful review of the facts and circumstances that existed in the circa 1918-1921 era provide compelling evidence that the 2022 final regulations reverse the priority ordering rule. Through its enactment of Section 901’s predecessor, Congress prioritized relief from double international income taxation over the revenue interest of the U.S. fisc when there was no consensus on international taxation norms at that time and countries were rapidly expanding their income taxation around the world. In contrast, instead of protecting against international double income taxation as its first priority, the 2022 final regulations protect the interest of the U.S. fisc as its first priority. As a general rule, the 2022 final regulations deny foreign tax relief to any foreign tax levy that fails to satisfy the attribution requirement in the regulations, which is a requirement that determines whether the foreign levy satisfies appropriate jurisdictional nexus and sourcing conformity requirements.77

When one juxtaposes the text, purpose, and policy goals that accompany the enactment of Section 901’s predecessor with the text, purpose, and policy goals of the 2022 regulatory changes, a remarkable contrast becomes evident. The following table sets forth the divergence of the text, purpose, and policy goals that guided the original enactment of the foreign tax credit as compared to the text, purpose, and policy goals of the 2022 final regulations.

76 See supra note 66 and accompanying text.
<table>
<thead>
<tr>
<th></th>
<th>Circa 1918-1921</th>
<th>2022 Final Regulations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Text</strong></td>
<td>The plain text of the statute eschews a nexus requirement. The original 1919 version had language that would mandate a nexus requirement, but it was affirmatively removed in 1921.</td>
<td>Notwithstanding the lack of textual authority, the plain text of the final regulations sets forth a jurisdictional nexus and sourcing conformity requirement.</td>
</tr>
<tr>
<td><strong>Purpose</strong></td>
<td>Prioritized avoidance of international double income taxation of U.S. persons over the revenue needs of the United States.</td>
<td>Prioritizes the revenue interest of the U.S. fisc over the interest of preventing U.S. double taxation of U.S. persons.</td>
</tr>
<tr>
<td><strong>Policy Goals</strong></td>
<td>Promoted international trade in the post-World War era in an effort to preserve global comity in advance of any actual agreement on international taxation norms. Such an agreement would only come decades later and was uncertain at the time of the credit’s enactment.</td>
<td>Protects the U.S. fisc by denying a foreign tax credit until an agreement on international taxation norms comes into existence, at which point the Treasury Department has indicated it would then revisit its eligibility requirements in light of a subsequent agreement on new norms.</td>
</tr>
</tbody>
</table>

In the end, the Treasury Department’s effort to impose a jurisdictional nexus and sourcing conformity requirement is contrary to the statutory changes (the text), the purpose, and the policy goals of the foreign tax credit as derived from the facts and circumstances that existed at the time of its original enactment. The foreign tax credit was a provision designed to restore peaceful trade before a broad agreement had been forged on international taxation norms. Instead of promoting the text, purpose, and policy goals of the statutory provision, the 2022 final regulations repudiate the historical text, purpose, and policy goals that prompted its original enactment.

At a more fundamental level, even if one were to believe that a jurisdictional nexus and sourcing conformity requirement should be imposed under Section 901, it is difficult to accept the Treasury Department’s assertion that a market jurisdiction’s taxation of remote sellers violates today’s international norms. As previously discussed, the OECD Inclusive Framework envisions that such taxation is allowed under its framework, and over 141 nations signed-on to that agreement. The United States also signed onto this agreement. Thus, as measured by international norms, taxation of remote sellers by market jurisdictions is consistent with the agreement forged by the OECD to which the Treasury Department has itself endorsed. The World Bank Group’s staff

---

78 See International collaboration to end tax avoidance, supra note 2.
79 See sources cited supra note 8.
recently stated that the development of Article 12B in the U.N. Model Treaty “provides an agreed international framework within which jurisdictions can adopt aligned approaches, leading to international consistency of treatment.”

Thus, it is becoming increasingly less plausible for the United States to contend that it is outside the norms of international consensus for market jurisdictions to assert taxation over digital activities conducted within their jurisdiction.

But even setting all of this aside, the U.S. Supreme Court has indicated that extraterritorial taxation over remote sellers is not inconsistent with U.S. jurisdictional norms in the state taxation case of *South Dakota v. Wayfair, Inc.* In that case, the state of South Dakota attempted to assert taxation over remote sellers that had no physical connection in the state. The taxpayer’s principle argument was that nonresident taxpayers could not be taxed in South Dakota if the nonresident did not have a substantial physical nexus within the taxing state. The Supreme Court upheld South Dakota’s right to tax remote sellers that had no physical presence. The Court, therefore, rejected the notion that the U.S. Constitution requires a physical nexus for a state to assert jurisdiction over remote economic activity connected to its taxing jurisdiction. The following statement indicates this rejection:

The physical presence rule has “been the target of criticism over many years from many quarters.” *Direct Marketing Assn. v. Brohl*, 814 F.3d 1129, 1148, 1150–1151 (C.A.10 2016) (Gorsuch, J., concurring). *Quill*, it has been said, was “premised on assumptions that are unfounded” and “riddled with internal inconsistencies.” Rothfeld, *Quill*: Confusing the Commerce Clause, 56 Tax Notes 487, 488 (1992). *Quill* created an inefficient “online sales tax loophole” that gives out-of-state businesses an advantage. A. Laffer & D. Arduin, Pro–Growth Tax Reform and E–Fairness 1, 4 (July 2013). And “while nexus rules are clearly necessary,” the Court “should focus on rules that are appropriate to the twenty-first century, not the nineteenth.” Hellerstein, Deconstructing the Debate Over State Taxation of Electronic Commerce, 13 Harv. J.L. & Tech. 549, 553 (2000). Each year, the physical presence rule becomes further removed from economic reality and results in significant revenue losses to the States. These critiques underscore that the physical presence rule, both as first formulated and as applied today, is an incorrect interpretation of the Commerce Clause.

---


The physical presence rule as defined and enforced in *Bellas Hess* and *Quill* is not just a technical legal problem—it is an extraordinary imposition by the Judiciary on States’ authority to collect taxes and perform critical public functions. Forty-one States, two Territories, and the District of Columbia now ask this Court to reject the test formulated in *Quill*. See Brief for Colorado et al. as Amici Curiae. *Quill*’s physical presence rule intrudes on States’ reasonable choices in enacting their tax systems. And that it allows remote sellers to escape an obligation to remit a lawful state tax is unfair and unjust. It is unfair and unjust to those competitors, both local and out of State, who must remit the tax; to the consumers who pay the tax; and to the States that seek fair enforcement of the sales tax, a tax many States for many years have considered an indispensable source for raising revenue.82

The fact that the vast majority of U.S. states joined in the suit and indicated that a virtual presence is a “sufficiently close connection” to a jurisdiction for it to assert taxation over those remote participants contradicts against the Treasury Department’s assertion in Treas. Reg. §1.901-2(b)(5) that a physical presence must exist in order for U.S. notions of nexus to exist. The Supreme Court further opined in *Wayfair* that any notion of a physical presence nexus requirement must give way to the “far-reaching systemic and structural changes in the economy” and “many other societal dimensions” caused by the Cyber Age.83 As a result, the Treasury Department’s assertion that a physical presence is required directly contradicts the Supreme Court’s statement that “[m]odern e-commerce does not align analytically with a test that relies on the sort of physical presence” and that “[t]his Court should not maintain a [physical presence] rule that ignores these substantial virtual connections to the State.”84

In response to this argument, the Treasury Department stated simply that the manner in which states determine their tax base has no bearing on the question for how the federal government should address that question.85 This manner of distinguishing *South Dakota v. Wayfair* fails to acknowledge that the Supreme Court articulated the contours of nexus in the U.S. sense and found that the assertion of virtual nexus is within the norms of what a U.S. taxing jurisdiction may impose. To restrict the reasoning applied in *Wayfair* to simply an expression of state taxation norms repudiates the public policy and economic commerce arguments that guided the Supreme Court’s decision.

---

82 Id. at 2092, 2095–96.
83 Id. at 2097.
84 Id. at 2095.
Thus, the Treasury Department’s position appears to be that an international norm cannot exist unless the U.S. Congress has enacted a comparable tax provision into U.S. law. That position is a decidedly nationalistic view of what constitutes an “international norm” and is dismissive of the Supreme Court’s own statements about jurisdictional nexus under the U.S. Constitution. Section 901 purports to provide a foreign tax credit to any foreign tax that is an income tax in the U.S. sense, and Congress early-on removed any language that would have augured for a jurisdictional nexus and sourcing conformity requirement. The U.S. Constitution does not restrict Congress’ jurisdictional nexus prerogatives in the manner contemplated by the 2022 final regulations, and a broad international consensus has reached agreement on extraterritorial taxation of remote sellers under the OECD’s Pillar One proposal. In light of all of the above, the Treasury Department’s position appears to be one that is a negotiating tactic and not one based on a principled position.

B. Does the Case Law Under Section 901 Support a Close Conformity Requirement as a Prerequisite to a Foreign Levy’s Eligibility for the Foreign Tax Credit?

Earlier in Part I.A., this article addressed the legislative history and the statutory text, purpose, and policy of the foreign tax credit to consider whether the Treasury Department’s regulations are consistent with those principles. However, the Treasury Department separately argued that an independent basis for its regulatory action is the judicial case law that has interpreted Section 901. In this Part I.B., this article considers whether the existing case law supports the 2022 regulatory amendments.

1. Biddle Doctrine and Early Cases

From the beginning, a central problem of the foreign tax credit has been to determine the contours of the foreign taxes for which it lies. The description of creditable taxes in Section 901(b)(1) is “any income . . . taxes paid . . . to any foreign country.” At the core of the complex statutory system governing the credit is a reference to foreign “income taxes,” a category later enlarged by the addition of foreign taxes “in lieu of” income taxes. Beyond its use of the two operative words, “income . . . taxes,” the Code says

---

86 I.R.C. § 901(b) (West 2022).
almost nothing more about creditable taxes. The bare terms do reveal, at a minimum, that the credit lies for “taxes” (rather than some other kind of payment) imposed on “income” (rather than some other base).

In deciding what foreign taxes represent income taxes eligible for U.S. foreign tax credit relief, two sentences of dictum in Biddle v. Commissioner have taken center stage. Biddle dealt with whether a U.S. shareholder of a British corporation subject to an integrated British corporate tax could be treated as having “paid” the taxes imposed on the corporation. In its discussion of the issue of whether taxes had actually been paid, the Court offered the following thought as to the analysis that should apply to determine whether the tax (if paid) was an income tax: “‘[I]ncome taxes paid,’ as used in our own revenue laws, has for most practical purposes a well understood meaning to be derived from an examination of the statutes which provide for the laying and collection of income taxes. It is that meaning which must be attributed to it . . . .” This assertion that U.S. principles guide the determination of whether or not a foreign levy is an income tax has been warmly embraced in subsequent cases that address whether a foreign levy is creditable as a foreign income tax.

The Biddle dictum states, rather than resolves, the problem. The notion of an “income tax” that emerges from U.S. principles is itself none too clear. The U.S. income tax is not a static creation. It evolves and is amended almost on an annual basis. Even if a foreign jurisdiction had a tax regime that was identical to the U.S. income tax at some point in time, it would be hard to imagine that it would remain so. In terms of applying U.S. principles to the rich variety of foreign levies to determine their status as an income tax, early cases and I.R.S. rulings had held

---

88 See generally I.R.C. § 901 (West 2022).
89 For a further discussion of the controversy regarding whether a payment was a tax applied on net income, see JOSEPH ISENBERGH & BRET WELLS, INTERNATIONAL TAXATION: U.S. TAXATION OF FOREIGN PERSONS AND FOREIGN INCOME ¶¶ 56.13–14 (Wolters Kluwer 6th ed. 2022).
90 302 U.S. 573, 579 (1938) (dictum).
91 See Isenbergh & Wells, supra note 89, ¶ 56.10.2.
92 Biddle v. Comm'r, 302 U.S. 573, 579 (1938)
93 Id. The context of this passage indicates that the Supreme Court was asking whether the British tax had been paid within the meaning of U.S. tax concepts, and not whether the tax was imposed on “income.” On the latter question, therefore, the statement is dictum (not a holding). Id.
94 Although Biddle dealt with whether U.S. or foreign law should be used to determine the identity of the technical taxpayer of the foreign tax, subsequent cases used the Supreme Court's statement that U.S. law, not foreign law, should broadly be used for purposes of applying the U.S. foreign tax credit rules including with respect to the question of whether a foreign levy was an income tax. See Comm'r v. Am. Metal Co., 221 F.2d 134 (2d Cir. 1955).
95 Some have called for a broader allowance of creditability of taxes beyond foreign income taxes. See Isenbergh, supra note 87, at 229.
that taxes levied on “imputed income” could be eligible for U.S. foreign tax credit relief if net income is what was attempted to be taxed and was so taxed.96 In general, these cases and early I.R.S. rulings took an expansive view of credit eligibility, allowing considerable latitude to the foreign country to define the manner in which a tax arrived at the net income it intended to tax.97 The case law did not accept the foreign law characterization of their own foreign levy. Instead, U.S. principles would determine whether the foreign levy was in substance an income tax. Thus, although the diversity of foreign taxes made the pre-1983 case law inconsistent at the outer edges, the substantive law was one that was based on a principle-based approach: if the foreign tax was designed to tax some amount of net income and predominantly did in fact tax some amount of net income in practice, then U.S. foreign tax credit relief was appropriate in order to prevent international double income taxation.98

2. **Bank of America** Standard

An important articulation for how to harmonize these decisions came in *Bank of America v. United States*.99 In the *Bank of America* case, the Court of Claims offered a cogent formulation of the Biddle standard for creditable income taxes.100 At issue in *Bank of America* was credit for foreign taxes imposed on the gross income of a U.S. bank.101 The Court of Claims declared that the standard for determining the foreign tax credit eligibility was as follows: “a direct income tax is creditable, even though imposed on gross income, if it is very highly likely, or was reasonably intended, always to reach

---

96 See Keen v. Comm’r, 15 B.T.A. 1243, 1246 (1929) (a tax on presumed income was calculated on nondomiciled persons who maintained a residence in France; income was presumed to be a minimum of seven times the rental value of their residence; held, French tax was an income tax eligible for U.S. foreign tax credit relief); see also Hatmaker v. Comm’r, 15 B.T.A. 1044, 1045 (1929) (same); see also Burke Brothers v. Comm’r, 20 B.T.A. 657, 660 (1930) (Indian tax on goat skins based on the difference between the average sales price of goat skins in their destination from the average sales price in Calcutta and reduced by certain transportation expenses; held, the presumptive tax was an income tax entitled to U.S. foreign tax credit relief); see also Rev. Rul. 53-272, 1953-2 C.B. 56 (Haitian tax imposed on business income computed by multiplying the rental value of the land and buildings by five and assessing an income tax on this imputed income; I.R.S. held this was an attempt to tax presumed income and was eligible for U.S. foreign tax credit relief); see also Rev. Rul. 56-658, 1956-2 C.B. 901 (Cuban tax on sugar mill operators assessed based on the amount of sugar produced times the average price for sugar and reduced by 60% for “deemed expenses;” held that this presumptive tax was creditable as an attempt to tax income).


98 See id.; see also ISENBERGH & WELLS, supra note 89, ¶ 56.10.


100 See id. at 518.

101 See id. at 514–15.
some net gain in the normal circumstances in which it applies . . . ." This standard, known as the *Bank of America* standard, operates within the constraints of *Biddle*, but it also countenances departures from strict conformity with U.S. tax norms. A tax can reach “some net gain” in the normal circumstances in which it applies without being strictly based on the same amount of net income as defined in the Internal Revenue Code.

As with the *Biddle* standard, the *Bank of America* standard is not simply a statement of the literal holding in the case. In fact, all the taxes actually under review in *Bank of America* were found *not* to be creditable. Those taxes, from several different countries, were imposed on the gross income of enterprises engaged in the active conduct of business. Because no deductions were allowed for the cost of producing the income, while the underlying business activity necessarily entailed expenses, the taxes were held non-creditable. The statement of principle, however—that a foreign tax is creditable if designed to reach *some* net gain—is generally favorable to taxpayers. Indeed, the government argued for a much stricter standard. The first case on creditable taxes after *Bank of America*, *Inland Steel Co. v. United States*, largely reasserts the same standard. *Inland Steel* concerns the creditability of an Ontario mining tax imposed on a tax base from which only operating expenses were allowed as deductions but not interest, depletion, royalty payments to private owners, or depreciation. The tax was imposed whether or not the ore produced by the taxpayer was sold. This tax was essentially similar to the one imposed in *Keasbey*, and the outcome was the same:

> When the mass of the omitted items in the [Ontario tax] are considered together and in combination as applied to plaintiff’s mining business, it is clear to us that tax does not seek to reach, or necessarily reach, any concept of net gain from the mining business which would be recognized as such in this country . . . . The exclusions are far too widespread and important to permit the conclusion that some net gain is sure to be reached.

---

102 *Id.* at 519-20.
103 See, e.g., *id.* at 524–25 (deciding that the foreign taxes levied on gross income from the taxpayer’s banking business were not creditable).
104 See *id.*
105 See *id.* at 524.
106 See *id.* at 523.
107 See *id.* at 523 n.20 (noting the I.R.S. position that any denial of cost recovery, even slight, should prevent creditability).
108 *Inland Steel Co. v. United States*, 677 F.2d 72, 80 (Ct. Cl. 1982).
110 *Inland Steel Co.*, 677 F.2d at 85. The same Ontario tax was found creditable in a
The Bank of America standard, as expressed in Bank of America and Inland Steel, thus looked at whether the foreign tax was likely to reach some amount of net gain. When that standard could not be met using U.S. principles, the foreign tax was not considered to be an income tax in the U.S. sense. The Tax Court also endorsed the Bank of America standard as the correct application of the Biddle doctrine. These cases obtained prominent attention in the preamble to the 1983 regulations because of their particular articulations of the Biddle doctrine.

3. The 1983 Final Regulations and The Grafting of Formal Requirements onto the Bank of America Standard

The regulations on creditable foreign taxes under Section 901 purport to be—and probably are—a comprehensive statement of doctrine. The oil crises of the 1970s forced a systematic scrutiny of whether payments to foreign countries were foreign taxes or were disguised oil royalties. In this regard, petroleum taxes often were at least in part determined on a formulary basis. See, e.g., Rev. Rul. 69-388, 1969-2 C.B. 154, revoked by Rev. Rul. 76-215, 1976-1 C.B. 194; Rev. Rul. 55-296, 1955-1 C.B. 386 (holding that a surtax imposed by Libya based on a posted price per barrel on holders of petroleum concessions was a creditable income tax), revoked by Rev. Rul. 78-63, 1978-1 C.B. 228; 69-552, 1968-2 C.B. 306 (holding that a surtax equal to a percentage of the posted price per barrel of oil was a creditable income tax). See T.D. 7918, 48 Fed. Reg. 46271, 46272 (Oct. 12, 1983).


Later case decided after the issuance of the 1983 regulations. See ISENBERGH & WELLS, supra note 89, ¶ 56.13.4 n.311.

The Tax Court endorsed and explained the Bank of America standard in Bank of America National Trust & Savings Ass’n v. Commissioner. In that case, the Tax Court explained the controlling standard in the following terms:

Perhaps the test which we and the Court of Claims have articulated will not provide that magic touchstone whereby every situation in this area can be precisely located in the spectrum of foreign taxes ranging from pure net income taxes on one end to pure excise, sales, or privilege taxes on the other. But we are convinced that the test is not ‘manufactured out of whole cloth,’ as petitioner would have us believe, and that it provides a rational and manageable basis for interpretation of section 901(b)(1).


regulations, the Treasury Department in 1983 issued a set of regulations on creditable foreign taxes under Sections 901 and 903. Much of the explosion of complexity is attributable to the Service's efforts, since 1983, to redefine the scope of the foreign tax credit eligibility rules to outright disallow the availability of foreign tax credits generated in “inappropriate transactions.” Although Congress' intent with respect to formulary taxes may be in doubt, there is no doubt that the Treasury Department wanted to overturn prior case law to the extent that prior case law had granted

---

115 On November 17, 1980, the Treasury Department issued temporary regulations that articulated formal criteria that a foreign tax would be eligible for U.S. foreign tax credit relief if and only if the foreign tax was equivalent to an income tax in the U.S. sense, and for this test to be met the foreign tax must meet three formalistic tests (the gross receipts test, the realization test, and the net income test). Treas. Reg. § 4.901-2(c) (1980). For an analysis of these temporary regulations and their impact on prior law, see Phillips Petroleum Co. v. Commissioner, 104 T.C. 256, 285 (1995). The effect of the 1980 regulations was that a levy paid by a petroleum company to a mineral-owning foreign government could be denied in its entirety if the effective tax rates for petroleum taxpayers were significantly higher than those imposed on nonpetroleum taxpayers. Treas. Reg. § 4.901-2(d) (1980). Prior case law had determined that foreign taxes represented income tax if they were “substantially equivalent” in nature to the U.S. income tax regime. See, e.g., N.Y. & Hond. Rosario Mining Co. v. Comm'r, 168 F.2d 745, 747 (2d Cir. 1948), rev'g and remanding 8 T.C. 1232 (1947). However, on April 5, 1983, the Treasury Department stated that a foreign levy would be eligible for U.S. foreign tax credit relief if and only if the “predominant character” of the foreign levy was that of an income tax in the U.S. sense, and these final regulations left the underlying formalistic three-pronged test for creditability essentially unchanged. See Notice of Proposed Rulemaking, 48 Fed. Reg. 14641 (Apr. 5, 1983); see also D. Kevin Dolan, General Standards of Creditability Under §§ 901 and 903 Final Regulations—New Words, Old Concepts, 13 TAX MGMT. INT'L J. 167, 168 (1984) [hereinafter New Worlds, Old Concepts] (“One can only guess whether there is any difference between those general standards [predominant character versus substantially equivalent standard in earlier case law] in terms of the degree to which foreign law must conform to U.S. tax principles”). These 1983 proposed regulations also set forth detailed guidance on dual capacity taxpayers that granted partial foreign tax credit relief if the foreign jurisdiction had a generally imposed income tax that applied outside the extraction industry. For an analysis of this regulatory evolution through the issuance of the 1983 proposed regulations, see Petroleum Income, supra note 114, at 7–8. The final regulations issued on October 12, 1983, softened this dual capacity standard by providing that partial foreign tax credit relief would be available for dual capacity taxpayers even if the foreign country did not have a generally applicable income tax that was imposed on non-extraction taxpayers. See id. at 3–6.


117 The prior 1983 final regulations effort to impose more formality than simply the Bank of America standard is addressed in further detail in Wells, Foreign Tax Credit War, supra note 71, at 1960–62.

118 For an example of a pre-1983 case that the 1983 final regulations intended to overrule, see Seatrain Lines Inc. v. Commissioner, 46 B.T.A. 1076 (1942), nonacq., 1942-2 C.B. 31. In Seatrain, Cuba had imposed a formulary tax upon realized gain. In order to resolve a dispute over the amount of deductible expenses, the Cuban government substituted a three percent tax on gross shipping income for a six percent tax on net profits. The Board of Tax Appeals held that the tax was creditable because the tax was imposed on gain realized under U.S. standards and because the intent of the lower gross tax was to simulate the earlier net income tax at that higher rate. Id. For a discussion of this more lenient line of authority, see OWENS, supra note 97, at 46. For an excellent summary of the prior case law and the efforts made in the 1983 final regulations to tighten up the standards for allowing foreign tax credit relief, see Petroleum Income, supra note 114, at 8; see also Coven, supra note 114, at 115–16.
foreign tax credit relief for a gross formulary tax that did not provide for sufficient cost recovery. Consequently, whereas the pre-1983 case law had utilized a substance over form holistic approach to determine whether a foreign levy was assessed on some amount of net income, commentators and the courts recognized that the Treasury Department’s final regulations represented an effort to impose stricter conformity in terms of the actual formal design of foreign law. To ensure nobody missed this conclusion, after issuing the 1983 final regulations, the I.R.S. revoked fifty years of prior Section 901 revenue ruling positions whenever those prior rulings were inconsistent with the government’s regulatory formalistic standard for credit eligibility. The additional changes made in 2022 simply heighten the reliance on the existence in the foreign tax law of formalistic

121 See New Worlds, Old Concepts, supra note 115, at 168–69 (“Fortunately, the regulations provide specific tests for determining whether the general Bank of America standard is satisfied.”). Mr. Dolan was in the government and played an active role in drafting the 1983 regulations. See Wells, Foreign Tax Credit War, supra note 71, at 1916 n.66.
122 See Texasgulf, Inc. v. Comm’r, 107 T.C. 51, 73 n.3 (1996), aff’d, 172 F.3d 209 (2d Cir. 1999) (quoting Dolan commentary cited in note 115 with approval). In discussing the import of the 1983 final regulations, the Tax Court observed as follows:
The preamble states that the regulations adopt the creditability criterion from certain cases to use in deciding whether the predominant character of a foreign tax is likely to reach net gain for purposes of section 1.901-2(a)(3)(i), Income Tax Regs. The preamble states that a tax is likely to reach net gain if it meets three tests provided in the regulations. The regulations provide objective and quantitative standards that were not used in cases which decided creditability of foreign taxes before the regulations became final. Regulations can supersede prior case law to the extent that they provide requirements and definitions not found in prior case law.
Id. at 68–69 (emphasis added); see also, e.g., Bowater Inc. v. Comm’r, 101 T.C. 207, 212 (1993); Nissho Iwai Am. Corp. v. Comm’r, 89 T.C. 765, 776–77 (1987).
123 Initially, the I.R.S. did not challenge the foreign tax credit generator aspects of foreign taxes paid under production-sharing agreements that generated inflated amounts of U.S. foreign tax credits. See Rev. Rul. 69-388, 1969-2 C.B. 154 (discussing how Indonesia imposed a special tax by contract for companies operating in oil and gas producing regions in Indonesia held to be a creditable “in lieu of” tax under Section 903; this ruling was colloquially known as “Indonesia I” in the industry). The I.R.S. subsequently revoked Indonesia I. See Rev. Rul. 76-215, 1976-1 C.B. 194, 10–13 (stating that the payment was, in substance, a royalty, not a tax, and therefore not eligible for U.S. tax credit relief under either Section 901 or Section 903; this ruling was colloquially known as “Indonesia II” in the industry). But, by the mid-1970s, the I.R.S. decided to launch an assault on these “disguised oil royalty arrangements” even as Congress added a new foreign tax credit basket to address this same phenomenon. See Coven, supra note 114, at 100–05 (analyzing reversal of the historic I.R.S. position as set forth in its prior rulings).
125 See Coven, supra note 114, at 101–03.
design hallmarks that conform to those found in the U.S. income
tax laws as a prerequisite for foreign tax credit relief.126

This was and is a fundamentally different approach than the
one utilized under the Bank of America standard, where the
determination was simply whether a foreign tax was designed to
impose a tax on some amount of net income. Thus, this effort to
impose formal design requirements into the eligibility rules adds
complexity and invites considerable dispute.127 In 1983, the
Treasury Department attempted to adopt the Bank of America
standard and graft onto it further formal design requirements at
the same time. Thus, on the one hand, in the midst of this 1983
revision of the Treasury Department’s regulations, the Treasury
Department lauded Bank of America and Inland Steel as
enlightened articulations for how U.S. principles should be
applied.128 The prior 1983 final regulations adopted the
“predominant character standard” that was articulated in the
Bank of America case, which in turn focused on whether or not a
foreign levy must reach some amount of net gain in the normal
circumstances in which it applied. But then, on the other hand,
the prior 1983 final regulations grafted onto this Bank of America
standard the additional requirement that a foreign levy also must
possess familiar design hallmarks of the U.S. tax system
(realization, gross receipts, and net income) in order to be eligible
for foreign tax credit relief.129

4. The 2022 Final Regulations Close Conformity Standard

In 2022, the Treasury Department amended its regulations to
impose even stricter formal conformity requirements in how the
foreign tax is designed. The heightened design formality imposed
in 2022 was motivated by a desire to deny foreign tax credit relief
for novel extraterritorial taxation regimes where the foreign
jurisdiction imposed a tax in a manner that does not closely
conform to how the United States determines nexus, income
sourcing, and/or does not closely conform to how the United States
determines the amount of gain subject to taxation.130 Thus, in an
effort to deny credit relief for novel extraterritorial taxes, the

127 At least one respected commentator, who was personally involved in the actual
drafting of the 1983 regulations, has expressed the view that the Service's efforts to apply and
amend the 1983 regulations go in a different direction than the decided case law or the
drafters of the regulations would have envisioned. See generally Kevin Dolan, Foreign Tax
Credit Generator Regs: The Purple People Eater Returns, 115 TAX NOTES 1155 (June 18, 2007).
Treasury Department now requires the foreign tax law to closely conform in its formal design structure to how the U.S. income tax laws are designed. This close conformity requirement is imposed under the regulations without any intervening statutory change to Section 901. The heightened close conformity requirement raises the question of whether the regulations fulfill the core policy goal of preventing double income taxation as envisioned by Section 901. The fundamental goal of the foreign tax credit regime is to mitigate the evils of international double taxation.\textsuperscript{131} And, the statutory provision contains no restriction in terms of credit eligibility by reason of the sourcing conventions or nexus principles utilized by a foreign jurisdiction.

What is more, the 2022 final regulations outright repudiate the \textit{Bank of America} standard that considered whether or not a foreign tax was assessed on some amount of net gain. In its place, the 2022 final regulations require the foreign tax to utilize a tax base to determine the amount of gain subject to tax that closely conforms to how the United States determines the amount of gain subject to taxation under U.S. tax law.\textsuperscript{132} The preamble to the 2022 final regulations only cites \textit{Bank of America} for the proposition that a foreign levy must provide cost recovery and not for its predominant character standard, and cites \textit{Inland Steel} for the proposition that U.S. law is applied to determine eligibility.\textsuperscript{133} At that point, the 2022 final regulations eschew the manner in which those cases articulated the application of U.S. principles in favor of a restrictive close conformity standard not found in the case law. The 2022 final regulations remove all references to the “predominant character standard” applied in “normal circumstances.”\textsuperscript{134} The case law that articulated the use of U.S. principles did so as part of its effort to determine whether or not a foreign law attempted to reach some net gain. That determination would be made using U.S. principles, not foreign law principles.\textsuperscript{135}

\textsuperscript{131} See \textit{supra} note 66 and accompanying text.
\textsuperscript{132} See Treas. Reg. § 1.901-2(b) (2022).
\textsuperscript{134} See Treas. Reg. § 1.901-2(b) (2022).
\textsuperscript{135} The government explained that its regulatory standard based the net gain standard articulated in existing case law, but then attempted to constrict that standard to a formalistic standard. See T.D. 7918, 48 Fed. Reg. 46272, 46273 (Oct. 12, 1983). After endorsing \textit{Bank of America} and \textit{Inland Steel} as authority for mandating that each foreign tax must separately and formalistically satisfy pre-defined formal design features of the gross receipts test in Reg. Section 1.901-2(b)(3)(i), a realization test in Reg. Section 1.901-2(b)(2)(i), and a net income test in former Reg. Section 1.901-2(b)(4)(i), the 1983 final regulations subsequently provided that each such test must be separately met in order for a foreign levy payment to qualify for U.S. foreign tax credit relief. \textit{Id}. The Treasury Department was transparent in its desire, stating in the preamble to T.D. 7918:
That was the inquiry. Whether the amount of net gain corresponds to the same amount of net gain that the U.S. tax laws would have calculated was a non sequitur. The government in Bank of America argued that a foreign jurisdiction must design its tax base to closely conform to the manner in which U.S. tax base is constructed in order to become eligible for a foreign tax credit, but that position was rejected in Bank of America.

The 2022 final regulations require a foreign tax to either satisfy a Platonic idyllic notion of an income tax where exactly the right amount of net gain is determined with all significant expenses allowed as a deduction over some time period, or else the foreign law must deviate away from the Platonic optimum income tax base in a manner that closely adheres to how the United States income tax laws deviate from that Platonic optimum.\textsuperscript{136} If the foreign jurisdiction's law deviates from the Platonic optimum income tax in a manner that closely conforms to the deviations found in the U.S. tax laws, then that foreign law must diligently be amended whenever U.S. tax laws are amended to retain that close conformity.

The regulations set forth three tests for determining if a foreign tax is likely to reach net gain: the realization test, the gross receipts test, and the net income test. All of these tests must be met in order for the predominant character of the foreign tax to be that of an income tax in the U.S. sense.

\textit{Id.} The government is adamant in its litigating positions that a foreign tax must meet the formalistic tests set forth in Reg. Section 1.901-2(b) to be eligible for U.S. foreign tax credit relief. See Opening Brief for Respondent at 95, PPL Corp. v. Comm’r, 135 T.C. 304 (2010) (No. 25393-07) 2009 WL 6946860; Reply Brief for Respondent at 98, PPL Corp. v. Comm’r, 135 T.C. 304 (2010) (No. 25393-07), 2009 WL 6946860 (“The regulation provides . . . specific tests, all of which a foreign tax must satisfy to be deemed an income tax in the U.S. sense, and therefore creditable. These regulatory tests neither permit nor require the application of these tests to the ‘substance’ of the tax.”); Reply Brief for Respondent at 38, Entergy Corp. v. Comm’r, 683 F.3d 233 (5th Cir. 2012) (No. 25132-06), 2008 WL 8070871 (“Finally, analysis of pre-regulation case law does not assist in the resolution of this case, since petitioners do not dispute that the U.K. Windfall Tax must satisfy all . . . of the net gain requirements of the regulations to qualify as a creditable tax.”). For additional information, see Opening Brief for Respondent at 95, PPL Corp. v. Commissioner, 135 T.C. 304 (2010) (No. 25393-07) 2009 WL 6946860, which provides:

If a foreign tax fails to satisfy the “net gain” requirement of the Regulations, it is not creditable for U.S. tax purposes. And the ‘net gain’ requirement requires an analysis of neither the underlying purpose of the foreign tax nor the components of the foreign tax (to determine, for instance, if the Profit-Making Value is a generally accepted method for valuing a Windfall Tax Company). Simply, the “net gain” requirement requires that a foreign tax satisfy each of the objective tests (realization, gross receipts, and net income) to be creditable. The U.K. Windfall Tax fails to satisfy the net gain tests, and therefore it is not a creditable tax.

\textit{Id.} In proposed regulations issued in 2020, the government again relied on this prior case law as the basis for the net gain requirement. See generally Notice of Proposed Rulemaking, REG-101657-20, 85 Fed. Reg. 72078, 72089 (Nov. 12, 2020). But then it asserted that a further tightening of the net gain requirements was appropriate. \textit{Id.} The government adopted this heightened formality in 2022 through significant amendments to those regulations at that time, thus in effect doubling down on the litigating position that it lost in \textit{PPL}. See T.D. 9959, 87 Fed. Reg. 276, 292–99 (Jan. 4, 2022).

\textsuperscript{136} Treas. Reg. § 1.901-2(b)(4)(i)(A), (C) (2022).
over time. The 2022 final regulations cite no U.S. judicial case for this new standard because there is none. In terms of determining whether a foreign tax represents the Platonic optimum income tax or deviates in a manner that closely accords with the U.S. deviations, the 2022 final regulations look solely to how foreign law is written without any consideration to the “normal circumstances” in which the tax actually operates.

In terms of defining the Platonic optimum income tax, Treas. Reg. Section 1.901-2 begins in a noncontroversial manner by stating that a foreign tax must meet the regulatory “net gain” requirement in order for the foreign tax to be eligible for U.S. foreign tax credit relief. This “net gain” phraseology harkens back to prior judicial case law that took a holistic, substance over form inquiry into whether the foreign levy was assessed on some amount of net gain. However, it is at this point where the 2022 final regulations eschew the flexibility allowed under prior case law, stating that a tax will be conclusively determined to not meet the net gain requirement unless the foreign tax levy satisfies four specific formal design features. Specifically, the foreign tax must satisfy the realization test, the gross receipts test, the cost recovery test, and the attribution (or jurisdictional nexus

---

138 See generally id.
139 The regulatory requirement that the practical impact of foreign tax levy must be determined based on a substance over form inquiry was widely interpreted as inviting the use of empirical evidence and other proof to determine the true nature of the foreign levy. See Former Treas. Reg. § 1.901-2(a)(2)(ii)(E)(3) in T.D. 7918, 48 Fed. Reg. 46272, 46277 (Oct. 12, 1983). Consistent with this approach, in Texasgulf, Inc. v. Commissioner, the Tax Court held that the Ontario Mining Tax at issue in Inland Steel was creditable because for the taxpayer and others, the processing allowance did in fact “effectively compensate” for the disallowance of other cost recovery. 107 T.C. 51, 70 (1996). It is important to note that the Tax Court in Texasgulf accepted empirical evidence to determine whether a processing allowance provided a recovery of all significant expenses for the industry as a whole. See id. This decision was subsequently affirmed. See Texasgulf, Inc. v. Comm’r, 107 T.C. 51, 69 (1996), aff’d 172 F.3d 209 (2d Cir. 1999). The government asserted that the elimination of any need to look at the “normal circumstances” in which a foreign tax operates in practice was justified for administrability reasons as it was a more objective standard. See T.D. 9959, 87 Fed. Reg. 276, 292 (Jan. 4, 2022); Notice of Proposed Rulemaking, REG-101657-20, 85 Fed. Reg. 72078, 72089 (Nov. 12, 2020).
142 Treas. Reg. § 1.901-2(b)(1) (2022). For an excellent summary of the prior case law and the efforts made in the 1983 final regulations to tighten up the standards for allowing foreign tax credit relief, see Petroleum Income, supra note 114, at 5–6.
and sourcing conformity) requirement.\textsuperscript{146} For present purposes, the cost recovery requirement is of particular interest.

The reader by now has divined that the cost recovery requirement is satisfied if significant costs are recovered. Taxes predicated entirely on gross receipts or gross income do not satisfy the regulatory cost recovery requirement unless “there are no significant costs and expenses attributable to the gross receipts included in the foreign tax base.”\textsuperscript{147}

In response to considerable comments, the Treasury Department issued proposed regulations in 2022 that relaxed its cost recovery requirement and provided additional safe harbors to partially address the strict conformity approach that it had promulgated in the 2022 final regulations.\textsuperscript{148} Under the proposed regulations, a foreign tax satisfies the cost recovery requirement if it permits a recovery of “substantially all” of each item of significant costs or expense,\textsuperscript{149} which is a more lenient standard to the one set forth in the 2022 final regulations that required that all significant costs must be recovered but still one that is significantly higher than the prior case law. The proposed regulations then state that whether a foreign tax permits a recovery of substantially all of each item of significant costs or expense is determined based solely on the terms of the foreign law.\textsuperscript{150} However, notwithstanding this statement in the preamble, it is difficult to see how the factual determination of whether or not “substantially all” of the significant expenditures

\begin{itemize}
  \item \textsuperscript{146} Treas. Reg. § 1.901-2(b)(5) (2022).
  \item \textsuperscript{147} Treas. Reg. § 1.901-2(b)(4)(i)(A) (2022).
  \item \textsuperscript{149} Prop. Treas. Reg. §1.901-2(b)(4)(i)(A), 87 Fed. Reg. 71271, 71280 (Nov. 22, 2022); Treas. Reg. § 1.901-2(b)(4)(i)(A) (2022). A comparison of the language in the later 2022 proposed regulations with the earlier 2022 final regulations is helpful to clarify the change: A foreign tax satisfies the cost recovery requirement if the base of the tax is computed by reducing gross receipts (as described in paragraph (b)(3) of this section) to permit recovery of the significant costs and expenses substantially all of each item of significant cost or expense (including capital expenditures each item of cost or expense related to the categories described in section (b)(4)(i)(B)(2) of this section) described in paragraph (b)(4)(i)(C) of this section attributable, under reasonable principles, to such gross receipts.
  \item \textsuperscript{150} Prop. Treas. Reg. § 1.901-2(b)(4)(i)(C)(1), 87 Fed. Reg. 71271, 71281 (Nov. 22, 2022). The Treasury Department indicated that the application of the substantially all test outside of the safe harbors would be based solely on the terms of foreign law and would not involve a review of empirical evidence. Notice of Proposed Rulemaking, REG-12096-22, 87 Fed. Reg. 71271, 71274 (Nov. 22, 2022). The proposed regulations indicate that whether or not an item of cost or expense, if not designated as a per se significant cost, is significant for purposes of the regulations is determined based on whether, for all taxpayers in the aggregate to which the foreign tax applies, the item of cost or expense constitutes a significant portion of the taxpayer’s total costs and expenses. See Prop. Treas. Reg. § 1.901-2(b)(4)(i)(B)(1), 87 Fed. Reg. 71271, 71280 (Nov. 22, 2022). This regulatory standard appears to mandate a review of how the tax applies in practice that necessarily requires one to analyze empirical evidence.
\end{itemize}
under all the relevant facts and circumstances were allowed could be made without looking into empirical evidence for how the foreign levy applied in practice.\textsuperscript{151} The safe harbors require one to determine whether the disallowances exceed some threshold of gross receipts or taxable income, and that requires an inspection into how the tax applies in practice in order to apply the safe harbors. Outside the safe harbors, notwithstanding the Treasury Department’s repeated assertions to the contrary in the preamble to its regulations, the regulatory provisions impose the requirement for a court to make a factual determination as to whether or not the foreign levy allowed substantially all of the significant costs as a deduction under all the facts and circumstances in which that particular tax applies.\textsuperscript{152} That determination logically predestines a court to inquire into whether the practical impact of a particular foreign law disallowance provision was in fact substantial in practice. That inquiry by necessity leads to review empirical evidence for how the foreign law applies in the actual facts and circumstances for the industry involved.

The dividing line between expenditures that are “significant” versus “insignificant” is a key decision point in terms of applying these regulations. The existing regulations provide helpful clarity on the treatment of flat taxes applied on the gross amount of fixed and determinable income.\textsuperscript{153} The existing regulations provide that cost recovery is not needed if the foreign gross basis levy applies to wage income or applies to investment income that is not derived from a trade or business.\textsuperscript{154} Under this standard, taxes on the gross amount of rents or royalties are on shakier ground, but at least some of them arguably are creditable as the regulations state that cost recovery is not needed if the costs are not significant.

\textsuperscript{151} The proposed regulations indicate that whether or not an item of cost or expense, if not designated as a per se significant cost, is significant for purposes of the regulations is determined based on whether, for all taxpayers in the aggregate to which the foreign tax applies, the item of cost or expense constitutes a significant portion of the taxpayers’ total costs and expenses. See Prop. Treas. Reg. § 1.901-2(b)(4)(i)(B)(1), 87 Fed. Reg. 71271, 71280 (Nov. 22, 2022). This regulatory standard appears to mandate a review of how the tax applies in practice that necessarily requires one to analyze empirical evidence.


\textsuperscript{153} The issue of whether Section 901 was intended to provide relief only for net income taxes or for gross income taxes has been the subject of scholarly debate for over sixty years, and there is little indication that the original Congress that adopted the U.S. foreign tax credit gave this issue much thought. See Surrey, supra note 68, at 819–22; see also H.R. REP. NO. 65-767, at 11–12 (1918); 56 Cong. Rec. 634, 667–78 (1918), http://www.govinfo.gov/content/pkg/GPO-CRECB-1918-pt12-v56/pdf/GPO-CRECB-1918-pt12-v56-4.pdf.

Moreover, even if a gross basis levy fails to satisfy the cost recovery requirement, these taxes may qualify as “in lieu of” taxes. A further exception also is provided for simplified presumptive tax regimes that apply only to “small business[es].” The small business exception now contained in the final regulations requires one to revisit prior guidance that had been given for presumptive tax regimes because the final regulations add a new legally relevant fact into the equation, namely is such a regime applied only to “small businesses.” The scope of this “small business” exception is unclear. The final regulations provide an example involving a simplified presumptive regime that applies to businesses that have gross revenue of less than $500,000 and declare that this satisfies the “small business” exception. Elsewhere in the U.S. tax law, a small business is defined as a business that has gross receipts of less than $25 million, but the regulations do not cross-reference this small business definition for purposes of applying its Section 901 regulations, nor do these regulations provide their own definition of a small business. Thus, the outer limits of when a business crosses over the threshold of a “small business” remain unclear.

More fundamentally, the introduction of a small business exception to the cost recovery requirement now creates a new dichotomy for how Section 901 is applied to taxpayers. Prior to the 2022 amendment to the final regulations, the eligibility for credit relief did not depend upon whether the taxpayer was large, small, or medium size. The same rules applied across the board to all foreign levies, regardless of the identity of the particular taxpayer group to which it applied. Moreover, the statutory provision does not indicate that the credit is dependent upon the size of the taxpayer’s business. Rather, Section 901(a) states that the provision applies to “the taxpayer,” and “the taxpayer” is defined

---

156 See I.R.C. § 903 (West 2022).
162 I.R.C. § 448(c) (West 2022).
expansively in Section 901(b). Thus, the regulations now differentiate among taxpayers even though the statute does not countenance such a distinction.

For income earned in an active trade or business, the existing regulations provide a heightened standard that requires the foreign jurisdiction to provide cost recovery in a manner that closely conforms to the manner in which cost recovery is allowed under U.S. income tax laws. In this regard, Treasury Regulation Section Reg. Section 1.901-2(b)(4) provides that the regulatory formulation of the cost recovery requirement is satisfied only if the tax allows (1) the recovery of all significant costs and expenses (including capital outlays) attributable under reasonable principles to gross receipts, or (2) the recovery of costs and expenses computed under a method that approximates or exceeds the amount of actual costs and expenses. These tests are alternatives. A tax that meets either one is therefore treated as satisfying the cost recovery requirement. Furthermore, although foreign law can allow for a different period for cost recovery than is allowed under U.S. law, the cost recovery requirement is not met if the deferral of cost recovery effectively represents a denial of such recovery. Taken as a whole, the final regulations posit that an income tax in the U.S. sense must allow for recovery of all significant business expenditures (or their economic equivalent) in some reasonable period unless there is a similar analogue to a disallowance provision under U.S. tax law. In addition, in the computation of the tax base, the foreign jurisdiction must utilize transfer pricing principles that comply with the arm’s length standard without taking destination-based criterion (like customer location) into account.

Given that a disallowance of any significant cost with respect to an active trade or business causes the foreign levy to fail as an income tax in its entirety, the determination of which costs are “significant” has profound importance. The regulations

---

163 See I.R.C. § 901(a)–(b) (West 2022) (defining taxpayers as citizens, domestic corporations, residents of Puerto Rico, certain nonresident aliens, and certain partnerships and estates without distinction as to size or shape).
165 The 2022 final regulations seemingly articulate an absolute standard that “all significant cost” must be recovered but the 2022 proposed regulations reduce this threshold to a requirement that “substantially all” of the significant cost must be recovered. See supra note 149.
168 See supra note 165 and accompanying text.
171 See supra note 165 and accompanying text.
provide that the significance of a particular cost or expenditure is determined based on whether it constitutes a significant portion of the total costs and expenses of all taxpayers subject to the tax. The existing regulations, however, then provide a list of “per se” significant costs and expenses. Included within that group of “per se” significant expenditures are the following: capital expenditures, interest, rents, royalties, wages or other payments for services, and research and experimentation expenditures. To determine whether a payment is made on debt or equity, foreign law (not U.S. law) is utilized. Thus, if a foreign country denies a deduction for a payment made on an instrument that under foreign law is treated as equity, the cost recovery requirement is met because a deduction for dividends is not a “significant cost.”

This outcome is the result even if the instrument is treated as debt for U.S. tax purposes and the associated payment is treated as interest (a per se significant cost) for U.S. tax purposes. Again, foreign law categorization of the nature of the expenditure controls for purposes of determining whether it is a significant cost.

However, after articulating a seemingly absolute standard that all significant costs must be recovered, the regulations then provide an exception for disallowance provisions that resemble disallowance provisions existing in the U.S. income tax laws. In this regard, a foreign tax is considered to permit the recovery of significant costs and expenses if the foreign tax law limits interest deductions (an otherwise designated “per se” significant cost) based on a regime that is similar to the disallowance regime set forth in Section 163(j). Moreover, a disallowance regime that disallows interest and royalty deductions in connection with hybrid transactions based on principles similar to those underlying Section 267A is also an acceptable variation. Finally, the disallowance of expenses based on public policy considerations similar to those articulated in Section 162 also represents an acceptable disallowance that does not run afoul of the cost recovery

---

174 See id.
175 See id.
177 This helpful clarification was stated in the preamble to the final regulations. See T.D. 9959, 2022-3 I.R.B. 352.
178 See supra note 165 and accompanying text.
180 See supra note 165 and accompanying text.
requirement. These examples represent a non-exhaustive list. In addition, the existing regulations permit the non-deductibility of provincial income taxes against the national tax. The final regulations state that a disallowance intended to limit base erosion or profit shifting represents an acceptable disallowance provision that does not run afoul of the cost recovery requirement, but then the regulations take pains to provide only examples of base protection measures that are found in existing U.S. tax law.

The proposed regulations issued later in 2022 further ameliorate this close conformity standard by requiring that the foreign disallowance provision only bear some family resemblance to a disallowance provision that exists in the U.S. income tax laws. In this regard, if the foreign law disallowance provision is based on a disallowance that bears a family resemblance to “any principle” for disallowance a deduction in the United States, then it is eligible for a principle-based exception to the cost recovery requirement. Thus, the Treasury Department has given tacit recognition that cost recovery is a flexible standard under an

---

183 Id.
184 Id. (noting the use of the term “for example” as illustrating the rule and not limiting the rule).
185 Id.
186 Id.
187 See id.; Prop. Treas. Reg. § 1.901-2(b)(4)(i)(F), 87 Fed. Reg. 71271, 71281–82 (Nov. 22, 2022). The evolution of the regulatory standard here is worth noting. The original 2022 final regulations required the disallowance provisions to be consistent with the principles of the existing United States Internal Revenue Code. See T.D. 9959, 87 Fed. Reg. 276, 338 (Jan. 4, 2022). This standard was modified in a technical correction to T.D. 9959 issued shortly thereafter that clarified the foreign law disallowance provision only needed to be consistent with “any principle underlying United States principles, including principles that seek to limit base erosion and profit shifting and public policy concerns. See T.D. 9959, 87 Fed. Reg. 45018, 45020 (July 27, 2022) (correcting T.D. 9959, 87 Fed. Reg. 276 (Jan. 4, 2022)). The proposed regulations move this foreign law disallowance provision exception to the significant cost recovery requirement out of Treas. Reg. § 1.901-2(b)(4)(i)(C)(1) and into a new Prop. Treas. Reg. § 1.901-2(b)(4)(i)(F), and then the proposed regulations further ameliorate the conformity requirement by stating that a disallowance of all or a portion of an item of significant cost or expense does not prevent a foreign tax from satisfying the significant cost recovery requirement if the foreign law disallowance is consistent with any principle underlying the disallowances required under the income tax provisions of the Internal Revenue Code, including the principles of limiting base erosion or profit shifting and addressing non-tax public policy concerns similar to those reflected in the Internal Revenue Code. See Prop. Treas. Reg. § 1.901-2(b)(4)(i)(F), 87 Fed. Reg. 71271, 71281 (Nov. 22, 2022). Moreover, the disallowance of expenses based on public policy considerations similar to those articulated in I.R.C. § 162 also represents an acceptable disallowance that satisfies the principle-based exception to the cost recovery requirement.
income tax, but the Treasury Department purports to require any departure from cost recovery to either bear a family resemblance to a corresponding U.S. departure or else the foreign levy must afford substantially all of the significant costs cost recovery in order to be creditable.

This close conformity requirement in the 2022 final regulations or the family resemblance test in the 2022 proposed regulations raises as many questions as they answer. The U.S. tax laws are not static, so the application of any conformity standard requires a year-by-year inquiry. Thus, a foreign tax that limits interest deductions in a manner similar to Section 163(j) or that disallows royalty deductions in a manner similar to Section 267A can deviate along those lines without running afoul of the “all significant costs” requirement because those deviations conform to how existing U.S. tax law deviates from the Platonic ideal income tax. However, even if a foreign jurisdiction did have a regime that conforms to U.S. tax laws at some point in time, Congress or the foreign jurisdiction may amend their respective tax laws at any time thereafter. If one just focused on the United States, one could imagine very different tax laws depending on the political party that is in power. Thus, even if foreign tax law correlated to U.S. tax law at some point in time, it is an open question whether it would remain so over time or at least would remain so over all periods of time. The regulations, if literally applied, would require an annual review of the current design of each country’s tax laws on an annual basis because the question is not simply did the foreign jurisdiction assert taxation over some amount of net gain but rather did the foreign jurisdiction design its tax base in a manner that closely conforms (or, under the proposed regulations, bears a family resemblance) to the cost recovery allowance found in the U.S. tax laws in that particular year. Perhaps due to this reality, the final regulations provide a safe harbor by stating that a foreign levy treated as an income tax under an applicable U.S. treaty qualifies as a “foreign income tax” if paid by a U.S. citizen or resident that elects the benefits under the treaty. Until further administrative guidance is provided, one would expect that the ultimate determination of whether a foreign jurisdiction’s cost recovery mechanisms sufficiently conforms to those of the United States would create significant uncertainty if a U.S. tax treaty is not separately applicable.

---

189 See supra note 165 and accompanying text.
190 Id.; Treas. Reg. § 1.901-2(b)(4)(iv)(C) (example 3) (2022) (sets forth an example of a thin capitalization regime that is based on section 163(j) as it existed prior to 2022).
Along these same lines, it is important to keep in mind that a key design challenge is for nations to protect their tax base against profit shifting and base erosion strategies of multinational enterprises. The G20 and the G8 have each expressed concern over how countries should prevent the artificial shifting of profits to low tax jurisdictions. The OECD has engaged in a multi-year study designed to provide recommendations on how countries should address this profit-shifting phenomenon (the so-called “base erosion and profit shifting” or “BEPS” project). Source countries are actively designing tax base defense mechanisms to supplement their income tax collection efforts. The introduction of such tax base protection measures creates uniquely complex U.S. foreign tax credit issues under the final regulations because the safe harbor afforded to base protection measures leaves significant unanswered questions. In this regard, the only disallowance provisions that are illustrated as satisfying this base protection safe harbor are those that bear a family resemblance to base protection measures found in the U.S. tax laws. It is unclear whether and to what extent this base protection safe harbor exception could extend to base protection regimes that do not contain analogous provisions found in the U.S. tax laws.

Prior to the addition of this newfound exception into the final regulations, the Treasury Department had ruled unfavorably with respect to base protection regimes adopted in other countries. In this regard, as an example, many Latin American countries historically have relied on alternative minimum asset tax regimes to backstop their broad-based general income tax regime.
countries have viewed asset tax regimes as necessary anti-abuse measures to protect against base erosion from aggressive inbound tax planning. Asset taxes generally range from 0.2 percent to 2 percent and indirectly represent a limit on thinly capitalized companies. Some form of asset tax has existed at some point in the tax laws of Argentina, Bolivia, Colombia, Costa Rica, Ecuador, Guatemala, Honduras, Mexico, Nicaragua, Peru, and Venezuela.

Prior to the Treasury Department's regulatory changes in 1983, a business asset tax enacted to complement a country's collection of its general income taxes would probably have been viewed as a creditable foreign tax under pre-1983 case law. In fact, the Argentine government adopted its business asset tax only after it received assurance from the International Monetary Fund ("IMF") that the Argentine asset tax would be creditable in the United States. The Argentine government was later surprised to find out that the IMF's assurances that the Argentine asset tax would be entitled to U.S. foreign tax credit relief were incorrect.

With the notable exception of the United States, a survey of existing worldwide tax treaties reveals a broad international consensus that asset tax regimes implemented as part of the overall general income taxes of a foreign country should be eligible for foreign tax credit relief under bilateral income tax treaties around the world.

199 See, e.g., Dictamen D.A.L. 55/99 [Opinion by the Tax Legal Advisory Department of Argentina] (June 25, 1999). The theory for an asset tax is that a business asset should generate at least a minimum level amount of income (a return on asset) over a reasonable period of time. If this is not the case and the business is continued, then the assumption must be that there is unreported income. See Bret Wells, Tax-Effective Methods to Finance Latin American Operations, 28 INT'L TAX J. 21, 22-23 (2002) [hereinafter Wells, Latin American Operations].


201 See, e.g., Wells, Latin American Operations, supra note 199, at 22-23.


204 See id.

205 This is recognized explicitly in many treaties. Convenio Entre El Reino De España Y La República Argentina Para Evitar La Doble Imposición Y Prevenir La Evasión Fiscal En Materia De Impuestos Sobre La Renta Y Sobre El Patrimonio [The Argentina-Spain Tax Agreement], Arg.-Spain, art. 2(3)(b) and art. 23(1), Mar. 11, 2013, 69 TNI 1128, Doc. 2013-6458; Agreement between the United Mexican States and the Republic of Chile for the Avoidance of Double Taxation and the Prevention of Fiscal Evasion with Respect to Taxes on Income and on Capital (with protocol), Mex.-Chile, art. 2(3)(b)(ii) and art. 23(1)(1), Apr. 17, 1998, 2484 U.N.T.S. 350; Convention Between the Republic of Venezuela and the Swiss Confederation for the Avoidance of Double Taxation with Respect to Taxes on Income and on Capital, Switz.-Venez., art. 2(3)(b)(ii) and art. 23, Dec. 23, 1997, 2235 U.N.T.S. 39782; Convention Between the Kingdom of Norway and the Republic of Venezuela for the
Even though out-of-step with international norms, the I.R.S. has ruled that the “separate levy rule” requires an asset tax to be separately tested and, at present, is adjudged to not be an income tax in the U.S. sense.\(^{206}\) Now that the final regulations allow space for reasonable base protection measures to prevent profit shifting, the question is whether these asset tax regimes might now be viewed as being “consistent with any principle underlying the disallowances required under the Internal Revenue Code.”\(^{207}\) In 2017, the United States adopted an alternative minimum tax with respect to base erosion payments through the enactment of Section 59A.\(^{208}\) Are the policy goals of Section 59A sufficiently similar to the goals of alternative minimum asset tax regimes employed in Latin America? Even if these regimes could satisfy the base protection safe harbor of the cost recovery test, these regimes still pose concerns under the gross receipts and realization tests. As of the writing of this treatise, Rev. Rul. 91-45 remains outstanding with the consequence that asset tax regimes are not eligible for U.S. foreign tax credit relief under existing published guidance even though such regimes are a base protection measure designed to protect the foreign jurisdiction’s income tax base.\(^{209}\) Given the broad international consensus that foreign tax credit relief should be available for alternative minimum assets taxes, the fundamental question is: what is the U.S. tax policy justification for this divergence from this international consensus?

Perhaps the most significant discontinuity with the regulations is the fact that the I.R.S. has ignored the cost recovery requirement in several rounds of guidance on innovative foreign formulary tax levies that were adopted as part of a foreign country’s income tax laws. In this regard, Mexico enacted a tax in 2008 called the impuesto empresarial a tasa única [single business tax rate] ("IETU") and repealed the IETU as of January 1, 2014.\(^{210}\) The main goal of this tax was to fight tax evasion with Mexico’s underground

---

\(^{206}\) See Rev. Rul. 91-45, 1991-2 C.B. 336. Admittedly, Rev. Rul. 91-45 would allow Section 901 relief to apply if the Mexican asset tax payments were refunded and regular income tax payments were later made, but this requires the foreign country to carefully craft its asset tax laws; other Latin American countries with similar asset taxes have not done so, and it is difficult to articulate why they should. See generally I.R.C. § 901 (West 2022).


\(^{208}\) See generally I.R.C. § 59A (West 2022).


economy by requiring companies that do a large amount of business in cash to pay a two percent tax (increased to three percent as of January 1, 2010) on the deposit of currency above MXN 25,000.\textsuperscript{211} The IETU’s explicit goal was to stop tax evasion, so the tax did not target compliant taxpayers.\textsuperscript{212} The IETU was creditable against federal Mexican income tax. Because this tax did not allow deductions, tax scholars\textsuperscript{213} and the tax practitioner community understood that this tax failed to meet the formalistic cost recovery requirement set forth in the regulations.\textsuperscript{214}

Instead of issuing a ruling that set forth this result, the I.R.S. issued Notice 2008-3.\textsuperscript{215} In this notice, the I.R.S. said that this tax needed “study” and that “the IRS will not challenge a taxpayer’s position that the IETU is an income tax that is eligible for a credit.”\textsuperscript{216} The I.R.S. allowed interim creditability for the IETU without providing any coherent rationale for how this tax satisfied the cost recovery standards set forth in Treas. Reg. Section 1.901-2(b). The reality was, and is, that the I.R.S. simply did not want to apply its own overly formalistic cost recovery requirement.\textsuperscript{217} However, to achieve this result the I.R.S. simply did not apply its own regulations.

In 2010, Puerto Rico imposed a formulary excise tax on multinational enterprises operating within its borders.\textsuperscript{218} Instead of faithfully applying its existing 1983 final regulations and then applying the completely “in lieu of” standard of Section 903, the government stated in Notice 2011-29 that the provisions of this excise tax were “novel.”\textsuperscript{219} Because this excise tax qualified as “novel,” the Service stated that “pending resolution of these issues, the IRS will not challenge a taxpayer’s position that the Excise Tax is a tax in lieu of an income tax.”\textsuperscript{220} Thus, again, without any coherent explanation, the I.R.S. stated that it would not challenge the foreign tax credit eligibility of this formulary tax even though it did not (and in this author’s opinion could not)\textsuperscript{221} articulate a coherent rationale for allowing credit relief within the framework.

\textsuperscript{211} See id.
\textsuperscript{212} See id.
\textsuperscript{214} See Randall Jackson, supra note 210, at 2.
\textsuperscript{221} Others have reached the same conclusion. See generally Martin A. Sullivan, \textit{Puerto Rico Shows Tax Policy at its Best and Worst}, 77 TAX NOTES INT’L 467, 469 (2015); Martin A. Sullivan, \textit{The Treasury’s Bailout of Puerto Rico}, 73 TAX NOTES INT’L 267 (2014).
of the regulations. As a consequence, the I.R.S. is developing a de facto administrative working law that is unmoored to its existing regulations.222 This result is made all the more surprising as the government has issued regulations that make the noncompliance of this Puerto Rican tax even more clear.223 As a concession to Puerto Rico, the Treasury Department did clarify in its final regulations that the effective date of its regulations would not apply to disallow the creditability of this Puerto Rico foreign tax levy on or before January 1, 2023.224 Thus, a limited transition rule was provided for that levy. In June 2022, Puerto Rico amended its tax laws to replace its Excise Tax with an elective 10.5 percent income tax regime that would apply on the sale of goods and services into Puerto Rico.225 Finally, in another notice, the Treasury Department indicated that it was aware that Puerto Rico had enacted legislation that would allow taxpayers to amend their existing tax decrees with the consequence that the taxpayers would no longer be subject to the elements of the Puerto Rico taxes that led the Treasury Department to conclude that those levies were noncreditable, but may subject the taxpayer to greater income taxation in Puerto Rico due to this voluntary renegotiation of the tax concession.226 In this latter notice, the Treasury Department indicated that it would not contend that any portion of the resulting tax payment to Puerto Rico was a non-compulsory payment even if the ultimate amount paid to Puerto Rico under the modified tax decree was higher than the amount that would have been paid by the taxpayer under the previously negotiated tax decree as long as the modified tax decree was entered into before December 31, 2022, and the ultimate tax liability under the modified tax decree remained less than the general income tax liability that would have been imposed in Puerto Rico under its generally applicable income tax laws that would have applied if there had been no concessionary tax decree at all.227

223 See Treas. Reg. § 1.901-2(b)(5) (2022) (asserting that a foreign levy that does not have an appropriate jurisdictional nexus is non-creditable).
227 See id. § 3. The government recognized that Reg. §1.901-2(o)(5)(iii)(A) provides that where a foreign tax law provides a taxpayer with options or elections in computing its liability for foreign income tax whereby a taxpayer’s foreign income tax liability may be permanently decreased in the aggregate over time, the taxpayer’s failure to use such options or elections results in a foreign payment in excess of the taxpayer’s liability for foreign income tax. However, the Treasury Department then stated that given Puerto Rico’s status as a U.S. territory and to aid in Puerto Rico’s transition away from its prior tax decrees that imposed noncreditable levies the Treasury Department would not treat any additional tax liability as a non-compulsory payment. See id.
In 2008, the United Kingdom imposed a fixed £30,000 levy on U.K. non-domiciliary taxpayers.\textsuperscript{228} In Rev. Rul. 2011-19, the Service allowed the foreign tax credit for this tax.\textsuperscript{229} However, to reach this coherent outcome, the Service made the assertion that this levy was likely to reach net income even though it was a fixed amount and did not provide any deductions. As the press had reported at the time, this ruling cannot be reconciled with the regulatory net gain standard in Treas. Reg. Section 1.901-2(b).\textsuperscript{230} Even worse, the Service has not even tried to articulate a coherent rationale for how to harmonize this allowance of foreign tax credit relief with the standards set forth in its regulatory regime. Rev. Rul. 2011-19 has not been withdrawn or superseded even though recent final regulations make its ambivalence even more inexplicable.\textsuperscript{231}

Thus, taxpayers face difficult challenges in determining whether and to what extent Treas. Reg. Section 1.901-2(b) has in fact succeeded in changing the holistic approach that was characteristic of the existing case law. There have been public statements by Treasury officials indicating that they recognize that the conformity requirement imposed by the 2022 final regulations can provide overly harsh outcomes, but no further official guidance has been issued.\textsuperscript{232} Nevertheless, given the government’s prior practice of not applying its own overly formalistic standards in particularly harsh situations, it is foreseeable that the government may issue additional administrative guidance that would simply not apply the regulations to particular fact patterns, and there already have been calls for the government to do so. But even so, this practice, if continued, would also raise the question of whether the regulations should be reformed.

5. PPL’s Impact on the Application of the Biddle Doctrine

Significant disagreement exists in terms of the ongoing precedential impact of the Supreme Court’s decision in \textit{PPL Corporation v. Commissioner}.\textsuperscript{233} As discussed earlier, the Bank of

\begin{flushleft}


\textsuperscript{231} See Treas. Reg. § 1.901-2(a)(4)(i)(A) (2022) (stating that a foreign tax whose base is gross receipts or gross income for which no reduction is allowed for costs and expenses under foreign tax law does not satisfy the cost recovery requirement, even if in practice there are few costs and expenses attributable to all or particular types of gross receipts included in the foreign tax base).


\end{flushleft}
America standard had interpreted the Biddle doctrine to mean that U.S. principles are applied to determine whether a foreign tax was assessed on some amount of net gain. Familiar hallmarks of the U.S. tax system, such as realization, cost recovery, and gross receipts, were helpful guideposts to determine whether the foreign levy reached net income in substance, but the ultimate question remained whether or not a foreign levy was assessed on net income in the normal circumstances in which it applied. The 1983 final regulations purported to graft onto the Bank of America standard the additional requirement that a foreign tax must also be formally designed with the hallmarks of a U.S. income tax as an independent prerequisite before one could be eligible for foreign tax credit relief.

The Supreme Court’s decision in PPL occurred after the 1983 formalistic requirements were added to the Section 901 regulations, so it was decided after the Treasury Department had endorsed the Bank of America standard but then had grafted onto the Bank of America standard additional formalistic design hallmarks that must also be separately satisfied. Thus, consideration of the continued relevance of the PPL case is important for at least three reasons: (i) the decision was a unanimous decision of the Supreme Court where the Supreme Court was asked to decide whether a foreign levy, in addition to satisfying the Bank of America standard, must also possess formalistic design hallmarks of U.S. tax laws as enumerated in the Treasury regulations, (ii) the Supreme Court rejected the government’s effort to apply the prior 1983 final regulations to disallow a foreign tax credit for a foreign tax that did assess taxation on some amount of net gain even though that foreign tax did not meet the formal design hallmarks set forth in the prior final regulations, and (iii) the government has attempted, through its 2022 final regulations, to graft onto the Bank of America standard even greater formal conformity requirements after the PPL decision. Thus, the continuing relevance, if any, of the PPL decision has important implications for determining to what extent the final regulations are able to deny foreign tax credit relief in a situation where the foreign levy reaches some amount of net gain but fails to comply with the heightened formalistic design hallmarks of realization, gross receipts, cost recovery, and attribution requirements set forth in the final regulations.

234 See discussion supra Part I.B.2.
235 See discussion supra Parts I.B.2, I.B.3.
236 See discussion supra Part I.B.3.
The decision in *PPL* involved a so-called windfall profits tax adopted by the United Kingdom.237 The U.K. windfall profits tax at issue in the *PPL* litigation provided for a one-time twenty-three percent formulary assessment tax on all privatized utility companies.238 This tax was assessed on the difference between a company’s profit-making value239 and the price for which the company was privatized.240

Under the case law that pre-dated the 1983 regulatory changes, the above-described U.K. windfall profits tax would have been eligible for U.S. foreign tax credit relief under the *Bank of America* standard. Earlier iterations of U.K. excess profits tax regimes considered in the pre-1983 period had been found to be creditable,241 and I.R.S. administrative practice stated that a wide range of analogous excess profits tax regimes met the eligibility standards set forth in the pre-1983 case law.242 The I.R.S. had even ruled that a tax levy imposed on average profits spanning multiple years, much like the U.K. windfall profits tax that was the subject of the *PPL* litigation, was entitled to U.S. foreign tax credit relief,243 but these cases and administrative pronouncements preceded the 1983 regulatory amendments to Treas. Reg. Section 1.901-2(b).244 Thus, the *PPL*

---

237 See *PPL Corp.*, 569 U.S. at 331.

238 Id. at 332.

239 For this purpose, “profit-making value” was defined as its average annual profit per day over an initial period that was generally a four-year period, and then this amount was multiplied by nine—which was chosen as a baseline “price-to-earnings ratio.” Though described as a tax on excess value, the tax actually had the economic effects of a tax on excess profits, since the calculation of “value in profits terms” was based on average net income over the four-year period, as opposed to an actual measure of value (which could have easily been established from market data); thus, from an economic point of view, the U.K. windfall profits tax was a tax on excess profits. See *Brief for Rosanne Altshuler et al. as Amici Curiae Supporting Petitioners at 8–9, PPL Corp. v. Comm’r, 569 U.S. 329 (2013) (No. 12-43).*

240 See *Entergy Corp. v. Comm’r*, 683 F.3d 233, 234 (5th Cir. 2012).


242 Rulings concluded that a tax imposed at variable rates was creditable. See *Rev. Rul. 68-318*, 1968-1 C.B. 342 (stating that Italian tax on profits in excess of six percent of capital was creditable); *Rev. Rul. 56-51*, 1956-1 C.B. 320 (stating that Cuban tax on profits in excess of 1/10th of estimated real worth of capital was creditable); *Rev. Rul. 74-435*, 1974-2 C.B. 204 (stating that Swiss Cantonal tax imposed at variable rates on multi-year profits was creditable).

243 See *Columbian Carbon Co.*, 25 B.T.A. at 473 (stating that the Service contested timing of accrual, but not creditability of U.K. tax based on average profits of three-year period preceding assessment year); see also *Rev. Rul. 69-446*, 1969-2 C.B. 150 (stating that Swiss National Defense Tax, which is imposed on average profits for the two years preceding the assessment year, is an income tax).

244 See *Texasgulf, Inc. v. Comm’r*, 107 T.C. 51, 69 (1996), *aff’d* 172 F.3d 209 (2d Cir. 1999). Also, the government was categorical to the Tax Court, stating that the pre-1983 case law was of “little consequence” and that the 1983 final regulations superseded prior
case is interesting precisely because the taxpayer substantively satisfied the *Bank of America* standard for U.S. foreign tax credit relief under the historic pre-1983 case law criteria (a conclusion the I.R.S. National Office appears to have accepted before the litigation, or at least did not refute).\textsuperscript{245} Even so, the U.K. windfall profits tax failed to comply with the formalistic design hallmarks that the Treasury Department’s 1983 amendments to Treas. Reg. Section 1.901-2 had grafted onto the *Bank of America* standard. Thus, the facts set forth in the *PPL* case squarely put in issue whether the prior 1983 final regulations could require formalistic design hallmarks of a U.S. income tax to be met as a precondition for foreign tax credit relief even in situations where the foreign tax reached some amount of net gain (thus satisfying the *Bank of America* standard).

The Tax Court held that the taxpayer was entitled to foreign tax credit relief, finding as a factual matter that the U.K. windfall profits tax was designed to reach net income and did in fact tax net income in all cases.\textsuperscript{246} On appeal, the Third Circuit reversed the Tax Court’s decision.\textsuperscript{247} In its appeal to the Third Circuit, the government asserted,\textsuperscript{248} and the Third Circuit accepted,\textsuperscript{249} that the U.K. windfall profits tax used a tax base greater than gross receipts and therefore failed the gross receipts test contained in

\textsuperscript{245} It is interesting to note at this point that the I.R.S. National Office appeared to have agreed that the pre-1983 case law was supportive of the taxpayer’s position even before the *PPL* litigation; however, after analyzing that favorable case law, the I.R.S. National Office then argued that the government had authority to change the standards for creditability in its final 1983 Treasury regulations and stated as follows: “analysis of pre-regulation case law does not assist in the resolution of this case, since Taxpayer does not dispute that the U.K. Windfall Tax must satisfy the net gain test of the regulations to qualify as a creditable tax.” I.R.S Tech. Adv. Mem. 200719011 (May 11, 2007).

\textsuperscript{246} The Tax Court stated as follows:

Parliament did, in fact, enact a tax that operated as an excess profits tax for the vast majority of the windfall tax companies. The design of the windfall tax formula made certain that the tax would, in fact, operate as an excess profits tax for the vast majority of the companies subject to it. \[\text{[}\] Because both the design and effect of the windfall tax was to tax an amount that, under U.S. tax principles, may be considered excess profits realized by the vast majority of the windfall tax companies, we find that it did, in fact, “reach net gain in the normal circumstances in which it [applied],” and, therefore, that its “predominant character” was “that of an income tax in the U.S. sense.”


\textsuperscript{249} *PPL Corp.*, 665 F.3d at 67–68.
the 1983 final regulations.\textsuperscript{250} As an additional ground for reversal, the government asserted,\textsuperscript{251} and the Third Circuit accepted,\textsuperscript{252} that the U.K. windfall profits tax also failed to satisfy the realization test set forth in the 1983 final regulations. Because these formalistic criteria were not satisfied, the Third Circuit found that the U.K. windfall profits tax failed two of the mandatory tests contained in the 1983 final regulations and therefore was ineligible for U.S. foreign tax credit relief.\textsuperscript{253}

The Third Circuit denied foreign tax credit relief to the taxpayer in \textit{PPL}, but it never contested the Tax Court's factual determination\textsuperscript{254} that the U.K. windfall profits tax actually achieved its intended operational purpose of taxing only net income.\textsuperscript{255} Instead, the Third Circuit held that the 1983 final regulations had grafted onto the \textit{Bank of America} standard additional formal design hallmarks that must independently be satisfied beyond simply satisfying the \textit{Bank of America} standard, stating as follows:

Because the regulation repeats the phrase “predominant character” throughout its definitions, both the Tax Court and PPL on appeal suggest that it applies a “predominant character standard” independent of the three requirements. That is incorrect. We must

\textsuperscript{250} Id. at 65 (“In our view, PPL’s formulation of the substance of the U.K. windfall [profits] tax is a bridge too far. No matter how many of PPL’s proposed simplifications we may accept, we return to a fundamental problem: the tax base cannot be initial-period profit alone unless we rewrite the tax rate. Under the Treasury Department’s regulation, we cannot do that.”); Opening Brief for the Appellant, \textit{supra} note 248, at 31–32 (“The windfall [profits] tax was then imposed on the difference between profit-making value and flotation value, and a tax on the value of property does not have the predominant character of an income tax in the U.S. sense. Thus, the tax base for the windfall [profits] tax was completely divorced from any traditional concept of gross receipts.”).

\textsuperscript{251} The government asserted the following in its opening brief to the Third Circuit:

It is well-established that under U.S. tax law, a tax on value or appreciation is not a tax on realized income (and thus does not have the predominant character of an income tax in the U.S. sense). \textit{See Cottage Sav. Ass'n}, 499 U.S. at 559; \textit{Schmitt}, 208 F.2d at 821 (stating that it “is hornbook law of taxation” that a property owner “is not subject to income taxation upon the annual increase in value” of the property). Nor was the windfall tax a tax upon previously realized income. The fact that a company’s profit-making value was determined by reference to past profits does not convert the windfall tax into a tax on those past profits. Indeed, a tax on income-producing property does not become an income tax simply because the property’s value is calculated for tax purposes by reference to the amount of income the property generates.


\textsuperscript{252} \textit{See} \textit{PPL Corp.}, 665 F.3d at 67 n.3.

\textsuperscript{253} \textit{See} \textit{id}.

\textsuperscript{254} The Tax Court made specific findings of fact indicating that it found that the legislative intent for the U.K. windfall profits tax was to assess a tax on excess profits and the Third Circuit nowhere contests these findings. \textit{See PPL Corp. v. Comm'r}, 135 T.C. 304, 339–40 (2010), rev’d, 665 F.3d 60 (3d Cir. 2011), rev’d, 569 U.S. 329 (2013).

\textsuperscript{255} \textit{PPL Corp. v. Comm'r}, 569 U.S. 329, 337 (2013) (noting that the Third Circuit explicitly discussed its concerns regarding the gross receipts and realization requirements).
assess whether a foreign tax satisfies each of the regulation’s three requirements “judged on the basis of its predominant character.” Treas. Reg. § 1.901–2(b)(1), (b)(2), (b)(3), (b)(4). We may not, however, simply ask whether the “predominant character” of a foreign tax is that of a U.S. income tax without addressing the requirements. The Court of Claims did essentially that in a pair of cases that predated the Treasury regulation governing our case. See Inland Steel Co. v. United States, 677 F.2d 72, 80 (Ct.Cl.1982) (per curiam); Bank of Am. Nat’l Trust & Sav. Ass’n v. United States, 459 F.2d 513, 519 (Ct.Cl.1972).256

In one sense, the PPL case represents an odd case for disallowing foreign tax credit relief because the Tax Court made a finding of fact that the U.K. windfall profits tax operated as a tax levied on net income257 and resulted in a levy of some amount of net profits in all cases.258 Yet, the Third Circuit held that the U.K. windfall profits tax was non-creditable for U.S. foreign tax credit purposes because the formal design of the U.K. windfall profits tax did not use notions of gross receipts or realization that complied with U.S. standards with the consequence that the U.K. windfall profits tax (as drafted) failed to comply with the strict formalistic design standards that Treas. Reg. Section 1.901-2(b) had grafted onto the Bank of America standard.259

The Fifth Circuit in Entergy Corp. v. Commissioner held that this same U.K. windfall profits tax was entitled to U.S. foreign tax credit relief, thus creating a split in the circuits.260 In its evaluation of the Third Circuit’s plain textual reading of the 1983 final regulations, the Fifth Circuit in Entergy stated that the Third Circuit’s denial of foreign tax credit relief exalted “form-over-substance.”261 The Supreme Court granted certiorari

256 See PPL Corp., 665 F.3d at 64 n.1.
257 The Third Circuit is silent on this point, but the Fifth Circuit makes the statement categorically as follows: “the tax only reached—and only could reach—utilities that realized a profit in the relevant period, calculating profit in the ordinary sense (e.g. by subtracting operating expenses associated with generating the utilities’ income). This satisfies the net income requirement.” See Entergy Corp. v. Comm’r, 683 F.3d 233, 236 (5th Cir. 2012).
258 See PPL Corp., 665 F.3d at 67 n.3.
260 Id. at 237. The Fifth Circuit explained its disagreement with the Third Circuit’s analysis as follows:

In fact, as the record indicates, each utility could only be subject to the Windfall Tax after making a profit exceeding approximately an 11% annual return on its initial flotation value, and the Windfall Tax liability increased linearly with additional profits past that point. Moreover, the Third Circuit opinion seems to overlook that a tax based on actual financial profits in the U.K. sense necessarily begins with gross receipts, as, again, the record here indicates. London Electricity’s profit for purpose of the Windfall Tax was calculated by computing gross receipts less operating expenses. The Windfall Tax was designed to reach
in *PPL Corp. v. Commissioner* to resolve the circuit split.\(^{262}\)

The Supreme Court *unanimously* held that the U.K. windfall profits tax was entitled to U.S. foreign tax credit relief, thus reversing the Third Circuit’s decision and affirming the Tax Court’s original decision.\(^{263}\) Instead of discussing how the 1983 final regulations had attempted to impose additional formalistic design hallmarks on top of the *Bank of America* standard, the Supreme Court attempted to harmonize the prior 1983 final regulations with the pre-1983 case law, stating that Treas. Reg. Section 1.901-2(b) “codifies longstanding doctrine dating back to *Biddle.*”\(^{264}\) The Court omitted any serious discussion of the government’s assertion that its formal regulatory requirements sought to bring “structure and clarity” not found in the earlier case law.\(^{265}\)

The Third Circuit held that it could not simply apply the *Bank of America* standard in isolation because the Treasury regulations grafted onto that standard imposed additional formal design requirements that must be met in form.\(^{266}\) The government, in its brief before the Supreme Court, argued that its regulations imposed additional formal design prerequisites that must be met in addition to the prior case law standards and that its regulations should be afforded deference, citing the Supreme Court decision in *Mayo Foundation for Medical Education & Research v. United States.*\(^{267}\) The Supreme Court applied *Biddle* for the purpose of

---


\(^{263}\) *See id.* at 343.

\(^{264}\) *See id.* at 334–35.

\(^{265}\) *See Brief for the Respondent,* supra note 248, at 33.

\(^{266}\) *PPL Corp. v. Comm’r,* 665 F.3d 60, 64 n.1 (3d Cir. 2011), rev’d, 569 U.S. 329 (2013).

\(^{267}\) *Compare PPL Corp.,* 569 U.S. at 340–341 (where the Court discusses portions of the government brief dealing with pre-1983 case law), *with* Brief for the Respondent, supra note 248, at 33–43 (where the government asserts that the formalistic test set forth in the 1983 Treasury regulations is entitled to deference under *Mayo Foundation for Medical Education & Research v. United States,* 562 U.S. 44, 47 (2011)). The government’s argument was more robust in its brief before the Third Circuit and the Fifth Circuit as the following excerpt from its briefs in those proceedings so indicates:

[T]he Tax Court was required to accord the regulation *Chevron* deference. See *Mayo Found. for Med. Educ. & Research v. United States,* 131 S. Ct. 704 (2011). Moreover, “[b]ecause §901’s exemption from taxation is ‘a privilege extended by legislative grace,’” the regulation had to be “strictly construed.” *Texasgulf, Inc. v. Commissioner,* 172 F.3d 209, 214 (2d Cir. 1999) (quoting *Inland Steel Co. v. United States,* 677 F.2d 72, 79 (Ct. CL. 1982)). Instead, the Tax Court paid only lip service to the regulation. Although it discussed the regulation in summarizing the relevant legal principles (*PPL Op.* 24–26), the court went on to
determining whether the U.K. windfall profits tax was assessed on net income and found that its formal design, which did not comply with the formal design hallmarks set forth in the prior 1983 final regulations, was not a fatal defect. The Supreme Court’s nuanced handling of the government’s regulatory deference argument is interesting. Here is what the Court stated:

The Commissioner argues that . . . U.S. courts must take the foreign tax rate as written and accept whatever tax base the foreign tax purports to adopt. Brief for Respondent 28. As a result, the Commissioner claims that the analysis begins and ends with the Labour government’s choice to characterize its tax base as the difference between “profit-making value” and flotation value. Such a rigid construction is unwarranted. It cannot be squared with the black-letter principle that “tax law deals in economic realities, not legal abstractions.” Commissioner v. Southwest Exploration Co., 350 U.S. 308, 315, 76 S.Ct. 395, 100 L.Ed. 347 (1956). Given the artificiality of the U.K.’s method of calculating purported “value,” we follow substance over form and recognize that the windfall tax is nothing more than a tax on actual profits above a threshold.

Thus, the Supreme Court simply dismissed the government’s regulatory deference argument as unwarranted because any reading of Treasury regulations to require a form over substance analysis could not be squared with the black-letter principle that “tax law[s] deal[] in economic realities, not legal abstractions.” The Supreme Court eschewed any effort by the Treasury regulations to impose formalistic design requirements. In doing so, the Court opined that substance over form principles must be applied to effectuate the statutory purpose of Section 901 and that its application of those principles compelled the Supreme Court to conclude “that the windfall [profits] tax is [best viewed as] nothing more than a tax on actual profits above a threshold.”

Thus, the court considered at length the historical background and purpose of the windfall tax and its effect on the companies subject to the tax. It made no effort whatsoever to explain whether the windfall tax met any of the three regulatory subtests, all of which had to be met for the tax to be creditable.

**Compare** Opening Brief for the Appellant at 23–24, Entergy Corp. v. Comm’r, 683 F.3d 233 (5th Cir. 2012) (No. 10-60988), with Opening Brief for the Appellant, supra note 248, at 21–22.

268 See PPL Corp., 569 U.S. at 343–44 (“The tax is based on net income, and the fact that the Labour government chose to characterize it as a tax on the difference between two values is not dispositive under Treasury Regulation § 1.901–2. Therefore, the tax is creditable under § 901.”).

269 Id. at 340–41.

270 See PPL Corp. v. Comm’r, 135 T.C. 304, 330 (2010), rev’d, 665 F.3d 60 (3d Cir. 2011), rev’d, 569 U.S. 329 (2013) (“Respondent argues that the 1983 regulations alone control the creditability of the windfall [profits] tax because those regulations subsume or supersede prior caselaw and ‘neither require nor permit inquiry into the purpose underlying the enactment of a foreign tax or the history of a foreign taxing statute.’”).

271 PPL Corp., 569 U.S. at 340.

272 See id. at 340–41.
The Supreme Court decision can be read as a full-throated endorsement of solely applying the *Bank of America* standard notwithstanding that the Treasury Department—through its prior interpretive regulations—had attempted to circumscribe the *Bank of America* standard by adding formal design hallmarks of a U.S. income tax as an additional substantive prerequisite. The Supreme Court looked at the formal design hallmarks as simply helpful indicia, but even so, the ultimately dispositive question remained simply whether the U.K. levy was assessed on some amount of net income. Thus, instead of giving dispositive significance to the added regulatory formalistic design hallmarks set forth in the Treasury Department final regulations articulated, the Supreme Court placed a heavy judicial gloss over the prior 1983 final regulations to harmonize them with “longstanding doctrine dating back to *Biddle*” when in fact, the 1983 final regulations attempted to impose formality to the foreign tax credit eligibility analysis not found in the prior case law.

After its defeat in *PPL*, the government has doubled down on its regulatory efforts. In the preamble to its 2022 final regulations, the Treasury Department attempted to distinguish and narrowly construe the continuing import of the *PPL* decision in the following manner:

> The Supreme Court in *PPL* was applying the predominant character test in the existing regulations and was not interpreting the statute. Because the final regulations modify the standard for determining whether a foreign levy is an income tax in the U.S. sense, the final regulations do not conflict with the *PPL* decision. Thus, the Treasury Department and the IRS disagree with the comments' contentions that the 2020 FTC proposed regulations have inappropriately shifted the inquiry away from the substance, or the substantive economic effect, of the foreign tax.

At best, this characterization of the *PPL* decision is controversial. The Supreme Court decision rejected an invitation to apply Treasury regulations in a manner that would deny foreign tax credit relief to a U.K. levy that in substance was a net income tax but had failed to comply with the formal design features set forth in the Treasury regulations. The *PPL* decision utilized a substance-based inquiry that harkens back to the *Bank of America* standard. The 2022 amendments add further formalism and rigidity, which is in the same genre as the form over substance prerequisites that the Supreme Court categorized as “unwarranted” in the 1983

---

273 See id. at 330, 334 (stating that the regulations codify “longstanding doctrine dating back to *Biddle*” and then use “substance over form” principles to resolve the case).

regulations. The Supreme Court's decision is clear, and it is a unanimous decision that applied a substance over form approach to the foreign tax credit eligibility determination.

Thus, one is left with an important interpretive issue in terms of applying Section 901. This interpretive issue has to do with what one makes of the methodology utilized by the Supreme Court to decide the manner of applying Section 901 versus the methodology utilized in the Treasury regulations for applying Section 901. The Treasury Department recognizes this divergence, but it rationalizes the two methodologies by stating that \textit{PPL} is best viewed as a historic case that interpreted prior regulations and thus has limited applicability going forward. In addition, after distinguishing \textit{PPL} in this manner, the Treasury Department has added even greater formality into its regulations.

In contrast, the \textit{Bank of America} standard utilizes a substance over form analysis that looks to how foreign law operates in practice. In addition, the Supreme Court's actual holding and reasoning in \textit{PPL} sought to position that decision within the historic rationale of Section 901 to reach a result that the government itself understood (and argued in its briefs) was inconsistent with its very own final regulations. Thus, the holding of the \textit{PPL} case was one that rejected a regulatory effort to circumscribe the prior judicial case law. Viewed in that light, the government's 2022 amendments seek to repudiate judicial case law in an even harsher manner. The issue can be succinctly stated as follows: would a court really deny a foreign tax credit for a foreign levy that in substance is assessed on net income but does not meet the formal design requirements set forth in the final regulations? If that result occurs, then that outcome would create the type of double international taxation that the Supreme Court has stated is antithetical to the policy goal that the statute was intended to effectuate.

A final comment about \textit{PPL} is in order, and it relates to the Supreme Court's own interpretation of the \textit{Biddle} doctrine in the course of its \textit{PPL} opinion. In a two-sentence statement in \textit{PPL}, the Supreme Court offered its own further formulation of the \textit{Biddle} doctrine in this statement:

Instead of the foreign government's characterization of the tax, the crucial inquiry is the tax's economic effect. See \textit{Biddle, supra, at 579, 58 S. Ct. 379, 82 L. Ed. 431} (inquiry is "whether [a tax] is the substantial equivalent of payment of the tax as those terms are used in our own statute"). In other words, foreign tax creditability depends

\begin{footnotes}
\footnote{275} \textit{PPL Corp.}, 569 U.S. at 340–41.
\footnote{276} See Foreign Tax Credit Guidance, \textit{supra} note 274.
\footnote{277} \textit{PPL Corp.}, 569 U.S. at 343.
\footnote{278} See \textit{supra} note 66 and accompanying text.
\end{footnotes}
on whether the tax, if enacted in the U.S., would be an income, war profits, or excess profits tax.\footnote{PPL Corp., 569 U.S. at 335.}

This formulation (or reformulation, as the case may be) of the \textit{Biddle} doctrine has several important touchstones. Again, the Treasury Department has stated that its regulations are an effort to apply the \textit{Biddle} doctrine, so this recent Supreme Court rearticulation of the \textit{Biddle} doctrine has profound significance in terms of determining whether the Treasury regulations are a faithful articulation of the \textit{Biddle} doctrine. The Supreme Court said that \textit{Biddle} requires one to determine the economic effect of a tax. This harkens back to the idea that one should look to empirical evidence to determine the actual operation of the foreign levy in practice. This translation of the \textit{Biddle} doctrine (namely, looking to the economic effect of the foreign levy) is diametrically opposite to one that looks solely to the formal design of a foreign levy. Also, the above second sentence applies the \textit{Biddle} doctrine by asking a hypothetical question: would the foreign levy be considered an income tax \textit{if enacted in the United States}? The U.S. principles are used to determine the economic substance of the foreign levy in the first sentence, and U.S. principles are considered to determine whether the United States could have enacted the foreign levy under its own income tax laws under the second sentence.

This rearticulation of the \textit{Biddle} doctrine, as set forth in the above two sentences in the \textit{PPL} decision, provides significantly more latitude in terms of adjudicating the creditability of foreign taxes than what the Treasury Department believes the Supreme Court meant in \textit{Biddle}. The Supreme Court’s rearticulation of the \textit{Biddle} doctrine in its \textit{PPL} decision is reconcilable with the \textit{Bank of America} standard but expands upon it. Particularly, in terms of the second of the above two sentences, this understanding of the \textit{Biddle} doctrine represents a negative harbinger with respect to the Treasury Department’s argument that all significant costs must be allowed as a deduction in order for the foreign levy to be an income tax in the U.S. sense. Said differently, if the ultimate legal question is whether or not a foreign levy (if enacted in the United States) would be within the income tax authority of the Congress to enact, then existing case law provides strong support for the position that Congress need not afford cost recovery for all significant expenses for a U.S. tax to pass muster under the Sixteenth Amendment.

For example, an important case that addresses the necessity for cost recovery is \textit{Indopco, Inc. v. Commissioner}.\footnote{See generally \textit{Indopco, Inc. v. Comm'r}, 503 U.S. 79 (1992).} In \textit{Indopco}, the government contended that the allowance of cost recovery for
expenses was simply a matter of legislative grace and not an essential design feature of an income tax in the U.S. sense.\footnote{See Brief for the Respondent at 30–31, Indepco, Inc. v. Comm'r, 503 U.S. 79 (1992) (No. 90-1278).} Here, the government urged the Supreme Court to not allow an immediate deduction for expenditures if those expenditures provided a future benefit, even when no separate and distinct asset was created that could allow for future cost recovery.\footnote{It is important to note how many times the government states that there are “many” instances where significant expenses are not allowed for recovery under the U.S. income tax laws as of 1992:} Indeed, the situation presented in this case provides a perfect example of the inadequacy of petitioner’s “separate and distinct asset” test. Petitioner does not challenge the findings of the Tax Court (Pet. App. 30a) and the court of appeals (Pet. App. 12a) that the takeover transaction resulted in permanent benefits for petitioner. Application of the test urged by petitioner—under which outlays may be deducted in one year even though the benefits of the expense are reaped for many years in the future—would result in a distortion of petitioner’s income. For this reason alone, petitioner’s test should be rejected.\footnote{The courts have recognized many types of capital expenses that do not create or enhance any specific asset. 1 B. Bittker & L. Lokken, supra, ¶ 20.4.1, at 20-68. Most relevant are the “changed corporate structure” cases discussed at pages 17-19, supra. In these cases, as then-Judge Blackmun noted in General Bancshares, 326 F.2d at 716, even when the reorganization expenses “have not resulted in the acquisition or increase of a corporate asset, [they are treated as capital charges and] are not, because of that fact, deductible as ordinary and necessary business expenses.” Similarly, in Holeproof Hosiery Co. v. Commissioner, 11 B.T.A. 547 (1928), which was cited in General Bancshares, the court observed that “[i]t can be argued, and not without merit, that no capital asset is acquired when attorneys’ fees are paid in connection with an increase in capitalization, but it does not follow that the payments are ordinary and necessary expenses of the year when made.” 11 B.T.A. at 556. The mere fact that a corporation’s structure is not a “separate and distinct asset” does not mean that expenses incurred to alter its structure for the permanent betterment of the corporation are not capital in nature. . . . There are many other examples of business expenditures that have long been recognized as capital in nature even though they do not create or enhance any specific asset. The cost of an educational program that qualifies the taxpayer to enter a new trade or business is a non-deductible capital expenditure. Id. (emphasis added).} In a strongly-worded and
The Foreign Tax Credit Redux

staunchly pro-government opinion, the Supreme Court stated that an income tax in the U.S. sense means gross income and that the allowance of deductions is purely a matter of legislative grace.284 The following extended excerpt from the Indopco case is relevant for understanding the nature of the U.S. income tax system as now understood and interpreted by the Supreme Court:

In exploring the relationship between deductions and capital expenditures, this Court has noted the “familiar rule” that “an income tax deduction is a matter of legislative grace and that the burden of clearly showing the right to the claimed deduction is on the taxpayer.” The notion that deductions are exceptions to the norm of capitalization finds support in various aspects of the Code. Deductions are specifically enumerated and thus are subject to disallowance in favor of capitalization. Nondeductible capital expenditures, by contrast, are not exhaustively enumerated in the Code; rather than providing a “complete list of nondeductible expenditures,” § 263 serves as a general means of distinguishing capital expenditures from current expenses. For these reasons, deductions are strictly construed and allowed only “as there is a clear provision therefor.”285

The Supreme Court’s decision in Indopco makes abundantly clear that the Court would not entertain criticism of Congress’s refusal to allow cost recovery for a significant business expenditure as Congress has the unquestioned “power to condition, limit, or deny deductions from gross income in order to arrive at the net that it chooses to tax.”286 Consequently, post-Indopco, the formalistic cost recovery requirement that requires that all287 significant costs must be allowed as a deduction is at variance with what the government argued and the Supreme Court held in Indopco. In 1992, the government told the Supreme Court in Indopco that there are “many . . . examples” under U.S. tax law of business-related expenditures that do not create deductible expenses and never


284 See Indopco, 503 U.S. at 1043.
285 Indopco, 503 U.S. at 84 (citations omitted).
287 See supra note 165 and accompanying text.
provide cost recovery. Yet when judging a foreign country’s tax
levy, the net gain standard in the final regulations mandates that
the foreign levy provide for recovery of all significant expenses in
order for it to be considered an income tax “in the U.S. sense.”
The insistence by Treas. Reg. Section 1.901-2(b) that all significant costs must be recoverable in the foreign country’s tax
regime is diametrically opposed to what the government asserted in
Indopco about our own income tax regime.

In addition, the Supreme Court decision in Wayfair is also
relevant in terms of one’s understanding of the Biddle doctrine as
interpreted by the Supreme Court in PPL. In Wayfair, the
Supreme Court stated that notions of physical nexus “must give
way to the ‘far-reaching systemic and structural changes in the
economy’ and ‘many other societal dimensions’” of the Cyber-
Age. The Supreme Court then went on to state that “th[is] Court
should not maintain a [physical presence] rule that ignores [these]
substantial virtual connections to the State.” In PPL, the
Supreme Court stated that the Biddle doctrine means that
“foreign tax creditability depends on whether the tax, if enacted in
the U.S., would be an income, war profits, or excess profits tax.”
Because the Supreme Court’s decision in Wayfair makes it clear
that the Supreme Court would uphold any effort by Congress to
assert jurisdictional nexus over remote participants that have
continuous and sustained engagement with the U.S. marketplace,
this rearticulation of the Biddle doctrine in the PPL decision calls
into question the restrictive jurisdictional nexus standard set forth
in the 2022 final regulations. The Treasury Department asserts
that its restrictive jurisdictional nexus conformity requirement is
based on an application of the Biddle doctrine, but the Treasury
Department’s interpretation of the meaning of Biddle contradicts
the Supreme Court’s explanation of the Biddle doctrine in PPL. As
a result, the Treasury Department’s reliance on Biddle as the
basis for its authority to issue the jurisdictional nexus conformity

289 This is the standard in the existing Treasury regulations that provide that in
order for a foreign levy to qualify for credit relief then all but an insignificant amount of
allowances must never be less than the amount of the significant cost to which they are a
substitute unless the foreign levy applies only to small businesses. See Treas. Reg. § 1.901-
2(b)(4)(i)(B) (2022). A per se list of significant costs is provided, but the test is ultimately a
facts and circumstances test; disallowance regimes that are analogous to the United States,
including base protection measures, do not cause a failure to comply with the cost recovery
290 See supra note 165 and accompanying text.
292 Id. at 2095.
requirement in its 2022 final regulations is undercut by the Supreme Court’s own rearticulation of the meaning of the *Biddle* doctrine in its *PPL* decision.

Thus, it is safe to say that the Treasury Department’s understanding of the *Biddle* doctrine is diametrically opposite of the Supreme Court’s own rearticulation of the *Biddle* doctrine in the *PPL* case. This disagreement is more than an academic exercise. The Treasury Department relies on the *Biddle* doctrine as the basis for its authority to issue its conformity requirements in its 2022 final regulations and the family resemblance test in the 2022 proposed regulations. Yet the Supreme Court’s handling of the *Biddle* doctrine, as rearticulated in *PPL*, countenances far more latitude in the foreign jurisdiction’s design of its tax laws than is afforded by either of these regulatory pronouncements. The disagreement reaches a crescendo when the regulatory formal design prerequisites are not satisfied, but the foreign levy does assert taxation only over some amount of net gain in practice. In that situation, is the foreign levy eligible for foreign tax credit relief, or does the formal design defect cause the foreign levy to be ineligible for foreign tax credit relief? The Supreme Court stated in *PPL* that such an application of Treasury regulations was “unwarranted” and would be contrary to the longstanding “blackletter principle that tax law deals in economic realities, not legal abstractions” with the consequence that the substance of foreign law and not its form applies to determine the economic effect of a foreign law for purposes of determining credit eligibility under Section 901.” The level of disagreement between the Treasury Department and the judiciary in terms of how to apply the case law interpretation of the statutory provision of Section 901 has never been greater. Until one or the other backs down, the ultimate determination of the eligibility of a particular jurisdiction’s foreign tax is likely to remain controversial.

297 The Supreme Court has partially answered the question as to the result where an agency issues a regulation that is contrary to an existing case, stating that a “prior judicial construction of a statute trumps an agency construction otherwise entitled to *Chevron* deference only if the prior court decision holds that its construction follows from the unambiguous terms of the statute and thus leaves no room for agency discretion.” See Nat’l Cable & Telecomms. Ass’n v. Brand X Internet Servs., 545 U.S. 967, 982 (2005). The Supreme Court subsequently extended its *Brand X* standard further by stating that prior case law can *remove any ambiguity* so that “there is no longer any different construction that is consistent with [existing case law] and available for adoption by the agency.” See United States v. Home Concrete & Supply LLC, 566 U.S. 478, 487 (2012). The statutory language in Section 901 explicitly grants US foreign tax credit relief for any foreign tax that is in substance an income tax and does so without any statutory requirement that some
C. Treaty Implications Arising From the 2022 Final Regulations

The jurisdictional nexus and sourcing conformity requirement creates the very real possibility that a foreign jurisdiction’s taxes may fail the requirements of Section 901 and yet the United States already has a bilateral tax treaty with that jurisdiction. So, the next logical question is whether this newfound attribution requirement curtails eligibility for foreign tax credit relief vis-à-vis treaty partners of the United States. The final regulations answer this question by stating that a foreign levy that is treated as an income tax under an applicable U.S. income tax treaty qualifies as a “foreign income tax” if paid by a U.S. citizen or resident that elects the benefits under the treaty. The final regulations then provide that because controlled foreign corporations (“CFCs”) are not treated as U.S. residents under U.S. income tax treaties, those entities (as residents of a third country) do not qualify for benefits under U.S. treaties. Thus, the final regulations clarify that taxes paid by a U.S. treaty partner to a third-country controlled foreign corporation are separate levies that must independently satisfy the attribution requirement of Section 901 or Section 903. However, if the foreign country has agreed under a treaty with another jurisdiction to apply a source rule consistent with the U.S. source rule, then that treaty provision’s sourcing rule would be relevant to determine whether or not the foreign jurisdiction applied a sourcing rule that conforms to the U.S. rule as a result of the treaty provision.

The broad assertion in the 2022 final regulations that a tax payment made to a controlled foreign corporation is ineligible for an indirect credit to the U.S. shareholder appears to be an overstatement. For example, the 2016 U.S. Model Tax Treaty

formal design hallmarks must also exist. This statute could thus be read as unambiguous on its face as countenanced by Brand X, but if not, then the unanimous PPL decision arguably removed any ambiguity as contemplated by Home Concrete. Thus, after the PPL decision, there is arguably no ambiguity left as to the question of whether a foreign levy can be denied foreign tax credit relief based on form over-substance regulatory requirements. However, because the Court did not explicitly address the Chevron deference implications of its PPL decision as part of that decision, the issue of whether or not the Supreme Court’s decision in PPL supplants the Treasury Department’s authority to issue later regulations that interpret the statute differently remains unsettled. See id. at 493–94 (Scalia, J., concurring) (addressing this interpretive ambiguity when a Court settles a question but not explicitly addressing whether the statute had an ambiguity that satisfies the “Step 1” Chevron deference determination).

299 This is made clear in the preamble to the final regulations. See T.D. 9959, 87 Fed. Reg. 276, 292 (Jan. 4, 2022). However, the regulations provide that if the source rule is changed under a treaty to which the CFC is entitled to rely upon, then the modified source rule would potentially be tested to determine if the jurisdictional sourcing requirement is met. See Treas. Reg. § 1.901-1(a)(1)(ii) (2022).
explicitly provides for deemed paid credits for income taxes paid by a controlled foreign corporation.301 Furthermore, the existing treaty with Finland contains the same provision as Art. 23(2) of the U.S. Model Treaty and then explicitly provides that a withholding tax at source is a covered tax under the treaty.302 The question of whether this categorical denial of treaty benefits to foreign taxes paid by CFCs represents a regulatory effort to override treaties is made all the more relevant because those same Treasury regulations now explicitly recognize that an independent treaty-based foreign tax credit is available in some instances that are unsupported by the standards of Section 901. The scope of this treaty-based foreign tax credit is thus ambiguous now that Section 901 is no longer the controlling standard.

Regardless of the ultimate scope afforded to this newfound independent treaty-based foreign tax credit (whether available to only U.S. persons, or whether it extends to deemed paid credits for income taxes paid by controlled foreign corporations), its existence provides a “solution” that reopen an old Pandora’s box. In the early 1980s, a significant law review article argued that the Treasury Department had negotiated tax treaties that afforded foreign tax credit relief for foreign levies that failed to satisfy the Section 901 requirements.303 The article then posited that this independent treaty-based foreign tax credit was unmoored to domestic law and thus raised serious normative policy concerns.304 After that article, the Treasury Department set about a multi-decade effort to ensure that U.S. foreign tax credit relief was not afforded under a U.S. tax treaty in a manner that was not consistent with the contours of Section 901. In fact, the allowance of a foreign tax credit under a U.S. tax treaty defers to domestic statutory provisions as the authorizing mechanism. In relevant part, the italicized portion of the following excerpt from the U.S. Model Treaty makes this point clear:

In accordance with the provisions and subject to the limitations of the law of the United States (as it may be amended from time to time without changing the general principle hereof), the United States

304 See id. at 2–3 (1982).
shall allow to a resident or citizen of the United States as a credit against the United States tax on income applicable to residents and citizens: . . .the income tax paid or accrued to ________ by or on behalf of such resident or citizen.\(^{305}\)

The technical explanation reinforces this point, stating that the eligibility for a foreign tax credit is a creature solely of domestic statutory law, stating as follows:

[Although the Convention provides for a foreign tax credit, the terms of the credit are determined by the provisions, at the time a credit is given, of the U.S. statutory credit. Therefore, the U.S. credit under the Convention is subject to the various limitations of U.S. law (see, e.g., Code sections 901-908). For example, the credit against U.S. tax generally is limited to the amount of U.S. tax due with respect to net foreign source income within the relevant foreign tax credit limitation category (see Code section 904(a) and (d)), and the dollar amount of the credit is determined in accordance with U.S. currency translation rules (see, e.g., Code section 986). Similarly, U.S. law applies to determine carryover periods for excess credits and other inter-year adjustments.\(^{306}\)]

Thus, the historic understanding since the early 1980s has been that a treaty jurisdiction’s tax is eligible for credit relief only to the extent allowed under domestic law.\(^{307}\) In U.S. treaty negotiations, the Treasury Department must itself satisfy that a covered tax was compliant with the contours of domestic law. Given this understanding, the holding period requirements of Section 901(k) or 901(l), the disallowance of otherwise eligible treaty-based credits by reason of Section 901(m), or the limitations applied to foreign tax credits under Section 904 overrode any usage of credits allowed under the treaty because the ability to claim or utilize a foreign tax credit is ultimately always dependent on domestic law. An independent treaty-based foreign tax credit was a non sequitur.

With the above historical context in mind, the 2022 final regulations significantly alter the understanding of foreign tax credit relief under U.S. treaties. The regulations now create a situation where a foreign tax credit exists, even though that foreign tax would fail to meet the eligibility requirements under Section 901. Thus, the Treasury Department has resurrected the notion that an independent treaty-based foreign tax credit exists apart from domestic law by reason of a bilateral tax treaty between the United States and the treaty jurisdiction. The allowance of a

\(^{305}\) U.S. Model Tax Treaty, supra note 301, art. 23(2) (emphasis added).


treaty-based foreign tax credit that is unmoored to Section 901 represents a significant departure from prior practice. It opens a Pandora’s box of questions as to whether and to what extent domestic law is overridden by a treaty-based credit when the treaty does not purport to restrict or deny the credit based in the same detail as domestic law. Would the later in time treaty represent the grant of an independent treaty-based foreign tax credit that is not constricted by the disallowance and eligibility rules of Section 901 or Section 904? This obsolete question posited by Professor Gann is a highly relevant question once again.308

The import of the 2022 final regulations, therefore, is that it applies one set of rules to foreign jurisdictions that have a tax treaty with the United States and a different set of rules to those jurisdictions that do not. Given that the United States has tax treaties with all developed nations and only a few developing nations, the impact of this bifurcated handling of U.S. domestic law is that developing nations will be held to a harsher and stricter standard than developed nations. Developing nations do not have the tax resources that the United States has at its disposal, so it is reasonable and unexceptional to believe that developing nations would adopt conventions and rules that attempt to provide greater administrability and that do not closely conform to how the U.S. has designed its income tax laws. Developing nations now face a Hobson’s choice: agree to an income tax treaty with the United States, redesign tax laws309 to closely conform to the design hallmarks of the U.S. income tax laws, or be denied credit relief on their foreign tax levies. It is unclear why the United States has an interest in disadvantaging its own multinational enterprises in terms of investing in developing nations, or why the United States should pressure developing nations to adopt principles that conform to a developed nation’s tax laws. Thus, another remarkable divergence is highlighted here. In 1918, the United States, in a great act of statesmanship, afforded foreign tax credit

308 See Gann, supra note 303 at 2.
309 The new conformity requirements raise concerns as to creditability in many developing nations worldwide and pose a concern that the largest Latin American economy, Brazil, would not have any taxes eligible for U.S. foreign tax credit relief. See Letter from Timothy McDonald, Chair, & Rick Minor, Vice President & Int’l Tax Couns., U.S. Council for Int’l Bus., to Lily Batchelder, Assistant Sec’y (Tax Policy), Jose E. Murillo, Deputy Assistant Sec’y for Int’l Tax Affairs, & Kevin Nichols, Int’l Tax Counsel, U.S. Dept of the Treasury (Mar. 24, 2022), http://uscib.org/uscib-content/uploads/2022/03/USCIB-FTC-Treas.03.24.2022.final_.pdf [http://perma.cc/SD83-2WVS]. But see Stephanie Soong Johnston & Alexander F. Peter, Brazil Drafting Law for OECD-Aligned Transfer Pricing Revamp, 106 TAX NOTES INT’L 410 (2022) (stating that in response, Brazil has announced that it will reform its transfer pricing conventions to align with the OECD framework).
relief on a unilateral basis.\textsuperscript{310} The import of the 2022 final regulations is to condition allowance of foreign tax credit relief either on the requirement that a developing nation enter into a U.S. tax treaty or redesign its tax laws to closely conform to the laws of the United States.

D. Developing Nations’ Interest in International Taxation

The OECD Inclusive Framework seeks to both ring-fence the revenue claims of developing nations and assert some level of minimum taxation over multinational enterprises.\textsuperscript{311} The big winners in this arrangement appear to be developed nations in the European Union, but other big winners appear to be multinational enterprises that now have the OECD arguing on their behalf against the claims of developing nations that otherwise would have asserted additional taxation over residual profits earned from digital sales into their market economies.\textsuperscript{312} In a letter to the OECD, the United Nations Committee of Economic, Social, and Cultural Rights argued that the OECD did not provide an equal voice to developing nations and then argued as follows:

This [OECD-sanctioned] solution will bring about only minimal benefits to developing countries. According to OECD’s own estimates, it will reallocate around USD 125 billion of profits to market jurisdictions. However, that amount represents only around USD 10 billion in tax revenue for the countries which as noted by the South Centre is “a minuscule amount, especially when the annual scale of corporate tax avoidance ranges from 100-307 billion.”

... We wish to express our concern that the Two Pillar solution, as it stands, would significantly undermine the revenue collection and taxing rights of low and middle-income countries. This in turn will affect the availability of resources to ensure the progressive realization of all economic, social and cultural rights, as well as of the right to development, as expeditiously and effectively as possible. This is more


worrisome during times of severe resource constraints caused by a cumulative negative impacts of the COVID-19 pandemic, previous fiscal adjustments due to high levels of indebtedness and additional need for public resources to respond to health and social protection requirements of the population.\textsuperscript{313}

Other economic reports also indicate that the OECD Inclusive Framework largely benefits developed nations over developing nations.\textsuperscript{314}

In contrast to the OECD approach, the commentary to the U.N. Model Treaty makes clear that the U.N. approach leaves the ultimate allocation of taxation as a matter of negotiation among the treaty jurisdictions, as indicated in the following statement:

\begin{quote}
[The new article [12B] simply represents an approach to allocating taxing rights between two jurisdictions — the market jurisdiction and residence jurisdiction — that both have a valid claim to tax the income. . . My clear view is it’s not a new taxing right. It’s as old as the hills [and] you see it at the state level in the U.S. The problem we have is that the residence state taxing rights are also legitimate, so you have to have an allocation of taxing rules by treaty to try to prevent double taxation.

. . .

Countries’ common practice of relinquishing their market-based taxing rights through bilateral treaties does not imply that those rights do not exist. It’s entirely legitimate in domestic law to tax based on presence in the market. You should try to be moderate, bearing in mind your situation. But then you have to negotiate at the international level about how much of that taxing right is preserved, and countries are more and more saying [they] want to preserve more of those taxing rights.\textsuperscript{315}
\end{quote}

As the ongoing debate ensues between developing and developed nations in terms of what allocation of taxation rights


\textsuperscript{314} See Julie McCarthy, \textit{A Bad Deal for Development: Assessing the Impacts of the New Inclusive Framework Tax Deal on Low- and Middle-Income Countries} 3 (Brookings Global, Working Paper No. 174, 2022); Letter from Alex Cobham, Tax Just. Network, to Pascal Saint-Amans, Org. for Econ. Co-operation & Dev. (Feb. 18, 2022) (on file with author); see also \textit{The Effect of the OECD’s Pillar 1 Proposal on Developing Countries - An Impact Assessment}, OXFAM 1 (Feb. 17, 2022), http://webassets.oxfamamerica.org/media/documents/Pillar_1_impact_assessment_v2_25JAN2022.pdf?gl=1*gw4w7Tga*ODU3NTkxMzgwLjE2NzNyNjMyMTC*._ga_R58YETD6XK*MTRY3MjI2MzIxNjI4LjEuMTY3MjI2MzMwNC41Ny4wLjA. [http://perma.cc/BCG9-5PD4] (stating “[w]e already know that the OECD’s Pillar 2 grants almost all revenue to a handful of rich countries, while leaving less than 3% for the poorest countries’ and then finding that Pillar 1 provides little more than a 3% digital service tax would, making it questionable whether developing nations should implement this more complicated arrangement).\textsuperscript{315}

should be afforded to market jurisdictions, the Treasury Department’s 2022 final regulations have added to that jurisdictional debate in a manner that creates far-reaching consequences. For example, although Brazilian income taxes have historically been considered to be creditable under general U.S. tax principles, the 2022 final regulations require that a jurisdiction’s income tax utilize the U.S. notion of the arm’s length standard. For administrative convenience, Brazilian income tax laws have long utilized a variety of fixed margin presumptions for purposes of applying a minimum income tax or for applying presumptive tax regimes; now, those longstanding aspects of Brazilian tax law raise concerns that the entirety of the Brazilian income tax is non-creditable under the 2022 final regulations. It is easy to understand why developing countries with less administrative resources would rely on simplifying assumptions in order to make their income tax laws administrable, versus how a developed nation with significant resources would administer its income tax laws. The Treasury Department’s expectation that all nations must apply the level of rigor that the United States utilizes in terms of applying the arm’s length standard imposes a heightened standard on developing nations without any expressed Congressional endorsement for such treatment.

Congress has not endorsed the usage of Section 901 as a bargaining chip among nations. In fact, the regulatory effort to impose conformity standards in terms of jurisdictional nexus and

---

317 See Treas. Reg. § 1.901-2(b)(5)(ii) (2022). Brazil is the largest South American economy that does not have a tax treaty in force with the United States. Bob Michel & Tatiana Palcão, Pillar 1 as a Ticket to a Fairer Taxation for Low- and Middle-Income Countries, 106 Tax Notes Int’l, 655, 658 (2022). Commentators have recognized that Brazil has avoided entering into a bilateral tax treaty with the United States to preserve its sovereignty as a source state to tax profits realized in Brazil by nonresident aliens. See id. It is this type of market jurisdiction that the OECD’s Pillar 1 proposal would seek to circumscribe. In contrast, the U.N. approach in its Model Treaty would afford significant deference toward it in terms of allowing it to continue to forgo its path for exercising taxation over nonresident persons that earn digital income in its jurisdiction. See U.S. Model Tax Treaty, supra note 301, art. 23.  
319 See Isabel Gottlieb, India Urges Focus on Developing Country Needs in Global Tax Deal, Bloomberg Daily Tax Report (July 14, 2022, 1:25 PM), http://news.bloombergtax.com/daily-tax-report-international/sitharaman-urges-focus-on-developing-country-needs-in-tax-deal (arguing for the need for developing nations to have flexibility to adopt simplified transfer pricing formulas and treaty-based minimum tax regimes for administrative reasons); see also Reuven Avi-Yonah & Yoram Margalioth, Taxation in Developing Countries: Some Recent Support and Challenges to the Conventional View, 27 VA. TAX REV. 1, 9 (2007) (noting academic literature recommending greater reliance on withholding taxes).
the requirement of a foreign jurisdiction to utilize transfer pricing standards that closely conform to those of the United States involves the United States in the formal design of a foreign jurisdiction’s tax laws at a granular level—which departs from the principal goal of mitigating against the evils of international double taxation. The effort to reorient Section 901 into a prescriptive provision designed to promote conformity calibrates that provision to achieve a goal that is different from the original goal of eliminating double income taxation of U.S. persons.

II. SECTION 901 SHOULD HAVE BEEN AMENDED TO DENY FOREIGN TAX CREDIT RELIEF TO TOP-UP TAXES IMPOSED UNDER PILLAR TWO

The OECD’s Pillar Two project is designed to ensure that large multinational enterprises pay a minimum level of tax regardless of where they are headquartered and regardless of the jurisdictions where they operate.320 The OECD’s Pillar Two project introduces the concept of Global Anti-Base Erosion rules (so-called “GloBE” rules) that implement a minimum tax through interlocking rules.321 The first-in-line top-up tax is a qualified domestic minimum tax which is defined to be a minimum tax included in the domestic law of a jurisdiction and that: (a) determines the excess profits of the constituent entities located in the jurisdiction (domestic excess profits) in a manner that is equivalent to the GloBE Rules and (b) operates to increase domestic tax liability with respect to domestic excess profits to the minimum rate for the jurisdiction.322 The next-in-line top-up tax is the income inclusion rule (IIR) which is a top-up tax applied by the owner of the constituent entity.323 The IIR effectively operates by requiring a parent entity (in most cases, the ultimate parent entity) to bring into account its share of the income of each constituent entity located in a low-tax jurisdiction and taxes that income up to the minimum rate (after crediting any covered taxes on that income).324 “The IIR imposes a top-up tax only on that


321 Id. at 14–16.


323 See id. art. 2.2.

324 See id. art. 2.1–2.3.
portion of the low tax income of a foreign constituent entity which is beneficially owned (directly or indirectly) by the member of the group that applies the IIR (the Parent entity).

The last-in-line top-up tax is called the undertaxed payments rule (UTPR) which is a top-up tax that seeks to impose a top-up tax on a constituent entity and the tax is not a qualified domestic minimum tax or imposed by an entity other than the owner of the constituent entity. The UTPR acts as a backstop that can be triggered into operation if and only if a sufficient IIR did not already assess the minimum tax. Thus, the GloBE rules set forth a pecking order: the QDMT is in the front-of-the-line, the IIR is in the middle position, and the UTPR is the last-in-line top-up tax. However, the QDMT, the IIR, and UTPR apply only after covered taxes are taken into account, and those include the taxes paid by the constituent entity and taxes paid by the owners of a constituent entity under a CFC tax regime.

Importantly, the OECD Model Rules explicitly exclude top-up taxes from the definition of a covered tax. Thus, the OECD Model Rules make clear that the imposition of a top-up tax (a QDMT, a qualifying IIR, or a qualifying UTPR) is excluded from the definition of a covered tax and is thus ineligible for consideration with respect to whether a minimum tax has been paid in order to avoid a circularity problem. Because these taxes, in effect, are denied foreign tax credit relief under the Model Rules, these top-up taxes operate more closely in design to an international alternative minimum tax that would take second-chair status to the assertion of taxing jurisdictions that impose taxation under either a CFC tax regime or under the regular income taxes of a particular jurisdiction. However, as a concession to allow the source jurisdiction to assert taxation first on low-taxed income, any top-up tax assessed under a Qualified Domestic Minimum Top-Up Tax excludes any tax imposed under a CFC tax regime for purposes of computing the amount of the

325 REPORT ON PILLAR TWO BLUEPRINT, supra note 320, at 112.
326 See GLOBAL ANTI-BASE EROSION MODEL RULES, supra note 322, at 2.4–2.6.
327 See REPORT ON THE PILLAR TWO BLUEPRINT, supra note 320, at 15.
328 See GLOBAL ANTI-BASE EROSION MODEL RULES, supra note 322, art. 4.2. A qualified domestic minimum tax applied by the jurisdiction of the constituent entity would appear to also be a covered tax because it is recorded on the financial statements of the constituent entity per Article 4.2.1(a) and is not excluded by Article 4.2.2.
329 See id. art. 4.2.2.
330 Id.
Qualified Domestic Minimum Top-Up Tax, thus affording it first priority status as to the right to tax low-taxed income of the particular QDMTT jurisdiction. 332

Importantly, this nuance does not exist under Section 901 or its regulations. In this regard, a foreign income tax paid under a foreign law inclusion regime (like a qualifying IIR) would be considered attributable to the income to which the top-up tax relates,333 and the residency-based tax would appear likely to satisfy the attribution requirement334 and the other requirements of the Treasury regulations.335 Moreover, a qualifying UTPR and QDMT asserted by the jurisdiction of a constituent entity and a qualifying IIR asserted by the residency jurisdiction of the constituent entity’s owner are likely to be eligible for foreign tax credit relief under the existing regulations.336 This outcome represents a normative mistake, at least with respect to a qualifying IIR and a qualifying UTPR.337 If a foreign tax credit were allowed for IIR and UTPR top-up taxes, then the imposition of these top-up taxes would reduce the amount of actual tax imposed under GILTI, Subpart F, and the new U.S. corporate alternative corporate minimum tax with the consequence that the amount of tax paid on the particular country income would be further reduced below the minimum tax threshold. This, in turn, creates a circularity problem. The reduction of covered taxes due to the allowance of foreign tax credit relief for top-up taxes would create the need for a further top-up tax that would then again be triggered to apply, and so on. This circularity problem would ultimately lead to top-up taxes taking a first-priority status over the covered taxes that should be given first priority under the OECD framework (the CFC tax regimes of GILTI, Subpart F, and

335 Because the tax is applied to excess profits determined using income tax principles, it is likely the other requirements of the net gain requirement will be satisfied. See Treas. Reg. § 1.901-2(b) (2022).
337 Conceptually, a qualified domestic minimum top-up tax imposed by a jurisdiction on the income earned in that jurisdiction is a tax that should be afforded foreign tax credit relief in the U.S. in order to recognize that jurisdiction’s first right to assert taxation over the income arising from its own jurisdiction. The OECD has recognized this priority in its recent guidance. See OECD, OECD/G20 BASE EROSION AND PROFIT SHIFTING PROJECT: TAX CHALLENGES ARISING FROM THE DIGITALISATION OF THE ECONOMY – ADMINISTRATIVE GUIDANCE ON THE GLOBAL ANTI-BASE EROSION MODEL RULES (PILLAR TWO) para. 118.30 (2023), http://www.oecd.org/tax/beps/agreed-administrative-guidance-for-the-pillar-two-globe-rules.pdf [http://perma.cc/E94M-GB34].
the new alternative corporate alternative minimum tax). The OECD has already expressed the view that top-up taxes should not supplant a CFC tax regime, nor should they supplant a minimum tax imposed by the United States on its jurisdiction, but existing U.S. law does not provide for a mechanism to exclude these top-up taxes because no conforming amendment has been made to Section 901 to address this design challenge. This is a mistake in existing U.S. law. The U.S. foreign credit regime should deny foreign tax credit relief to all top-up taxes. This is made all the more urgent because it appears that several countries are moving forward with implementing Pillar Two.  

As a result, even though the Treasury Department signed on to the OECD Framework along with 137 other nations in October 2021, the fact remains that the 2022 amendments to the Treasury regulations, the corporate alternative minimum tax legislation enacted in 2022, and the 2022 Greenbook proposal identify the need to amend Section 901 to deny foreign tax credit relief for top-up taxes imposed under the auspices of the OECD Pillar Two project. For the reasons already addressed in Part I.C., an amendment to Section 901 that denies foreign tax credit relief to IIR and UTPR top-up taxes is not a treaty override because the allowance of a credit under U.S. tax treaties is made subject to the conditions of domestic U.S. tax law. Thus, the treaties defer to domestic law to define the terms of what taxes are eligible for tax credit relief. Thus, Congress could and should unilaterally deny U.S. foreign tax credit relief for any qualifying IIR and qualifying UTPR in order to prevent the imposition of those top-up taxes from reducing otherwise applicable U.S. taxation over that income.

CONCLUSION

The United States missed the correct turn and took a wrong turn with respect to the U.S. foreign tax credit implications of the OECD inclusive framework and the novel taxes that are being considered by other nations.

338 See Amanda Athanasiou, Yielding to Stakeholder Pressure, U.K. Delays Pillar 2 Implementation, 106 TAX NOTES INT’L 1585 (2022) (announcing delay but expecting implementation beginning in 2024). It is also believed that Canada, France, Germany, Italy, and Japan will also proceed to implement Pillar Two as well. See Reuven S. Avi-Yonah & Bret Wells, Pillar 2 and the Corporate AMT, 107 TAX NOTES 693 (2022).
The Foreign Tax Credit Redux

The wrong turn that was taken was to amend the Treasury regulations to impose a U.S. conformity and a jurisdictional nexus requirement. The Treasury Department, via the issuance of its 2022 final regulations, repudiated the text, purpose, and policy grounds that undergird the foreign tax credit since its adoption by restricting its scope in ways that eviscerate the intended goals of the foreign tax credit regime. These 2022 final regulations also represent a strong repudiation of the Supreme Court’s own rearticulation of the Biddle doctrine in the PPL decision by attempting to formulate an interpretation of the Biddle doctrine that is inconsistent with the Supreme Court’s interpretation of its own doctrine. The timing of these regulations is ironic. The world today has many similarities to the circa 1918-1921 era, albeit the drivers that create the similarities are different. The foreign tax credit was enacted to prioritize elimination of international taxation in the midst of the post-World War I chaos. In the circa 1918-1921 era, the United States, in statesmanlike fashion, took the unilateral step of mitigating against instances of international double income taxation during the period when no agreed international norms existed. The crushing war debts after World War I resulted in instances of international double income taxation when there was no consensus on norms. The United States afforded a foreign tax credit without any prerequisite agreement on international norms, and then in the next fifteen years, spearheaded an effort to forge international norms. The U.S. representative who was the architect of the foreign tax credit, T.S. Adams, spearheaded this effort to forge international taxation norms until his passing at which point Mitchell Carroll took over that role for the United States.342 The COVID-19 pandemic, not World War I, has created enormous strains on fiscal resources in this era. The explosive growth of the internet has allowed multinational enterprises to maintain a significant virtual presence with customers in market jurisdictions. There is an unlevel playing field between traditional brick-and-mortar businesses subject to jurisdictional taxation and virtual businesses that escape income taxation in those local market economies.343 Thus, for different reasons, the world is now again in a situation where internal norms of taxation are in the midst of reformulation. The OECD and over

342 See generally Carroll, supra note 59.
343 The commentary to Article 12B of the U.N. Model Treaty makes this point in the following manner:

In this regard, modern methods for the delivery of services allow non-residents to render substantial services for customers in the other country with little or no presence in that country. This ability to derive income from a country with little or no physical presence there is considered by the Committee to justify source taxation of income from automated digital services.

See U.N. MODEL DOUBLE TAXATION CONVENTION, supra note 11, at 4.
135 participating countries and jurisdictions all agree that reformulation is needed. The OECD has announced an extremely accelerated timeframe for implementing a new international consensus. It took more than a decade for an international consensus to be forged in the post-World War I era, so the OECD’s timeframe by that standard is ambitious.

Yet, although remarkable similarities exist in the two eras, the U.S. Treasury Department has forged a diametrically opposite policy approach in this era compared to the one that Congress chose in the circa 1918-1921 era. In 1918, Congress adopted a unilateral foreign tax credit before a consensus on international taxation norms was forged, and the United States worked for a consensus on international norms in the succeeding years. In contrast, in 2022, the Treasury Department has sought to deny foreign tax credit relief on destination-based taxes until a further international consensus on taxation of the digital economy is fully implemented. Congress in 1918 prioritized mitigation of international double taxation above the interests of the U.S. fisc and then worked to create a consensus on international taxation. In contrast, in 2022, the Treasury Department prioritized the interest of the U.S. fisc over the consequences of international double income taxation. Seen in light of its historical objectives and historical context, the 2022 final regulations eviscerate the text, purpose, and policy goals that guided the enactment of the foreign tax credit regime. It remains to be seen what a court will decide in terms of this reformulation, but in this author’s mind, the Supreme Court’s rearticulation of the Biddle doctrine in PPL provides a negative harbinger for the Treasury Department’s attempt to further impose formalistic design hallmarks onto the Bank of America standard.

Although the path taken by the Treasury Department represents a “wrong turn,” it is true to say that an adjustment was needed to Section 901 as part of the international agreement to implement the OECD Pillar Two recommendations. The “right turn” that should have been taken is that Section 901 should have been amended to deny foreign tax credit relief for top-up taxes, modeled after an income inclusion rule or an under-taxed payment rule. This is made all the more critical because Congress adopted a corporate alternative minimum tax that does not comply with the GloBE rules. By enacting a corporate alternative minimum tax that does not fit neatly with the GloBE rules, the newly enacted U.S. corporate minimum tax may represent a better outcome than if the United States had enacted a qualified IIR in compliance with the GloBE rules. But the enacted legislation contains a deficiency. What should have been done concurrently with the enactment of
this corporate alternative minimum tax (but was not done) was to make a companion amendment to Section 901 so that Section 901 would not afford foreign tax credit relief for any non-covered tax that is designated as a qualifying IIR or a qualifying UTPR under the GloBE rules. Failing to do so has put the residual U.S. tax jurisdiction at risk of being eroded through minimum taxes imposed by other nations in preference to the corporate alternative minimum tax imposed by the United States. The OECD framework envisions that top-up taxes modeled after the GloBE rules would not be afforded foreign tax credit relief among nations, so the United States’ denial of foreign tax credit relief for top-up taxes would have been consistent with the OECD proposal for how these top-up taxes should be handled under the OECD inclusive framework. The U.S.’ failure to make this conforming amendment to Section 901 represents a self-inflicted wound. Congress should correct this mistake by amending Section 901 to make it clear that top-up taxes under a qualifying IIR or a qualifying UTPR would not be afforded U.S. foreign tax credit relief. Doing so would ensure that the U.S. tax base is not eroded and that these taxes truly represent incremental “top-up” taxes that cause a multinational enterprise to be subject to the minimum tax rate. Thus, reform along these lines effectuates the policy goals sought by the OECD framework and also protects the U.S. tax base. It is now time for Congress to address this deficiency before it creates an inappropriate reduction of the U.S. tax base.
Chapman Law Review Debate: Does Originalism Work?

Kurt Eggert* and Lee Strang,† moderated by Tom Campbell‡

PROFESSOR CAMPBELL:

Good afternoon. We’re very happy to put together this debate stemming from Professor Eggert’s paper, *Originalism Is Not What It Used to Be*, published by the *Chapman Law Review*.¹ We are looking forward to a presentation by Professor Eggert, and I wanted to give a brief introduction of him, and then a response followed by Professor Strang. For background, I think all of us here know Professor Eggert, but I did the research, so I’m going to share it with you.

Professor Eggert is, of course, the director of the Alona Cortese Elder Law Clinic here at Chapman Law School, and full Professor of Law. He has his J.D. from UC Berkeley. He has given Congressional testimony and published on consumer protection, mortgages, gambling regulation, and, of course, elder law, and was a member of the Federal Reserve Board’s Consumer Advisory Council. Professor Eggert is the author of the piece which is up in this debate.

Professor Lee Strang, most graciously, has come across the continent to be with us today. He is the John W. Stoepler Professor of Law and Values at the University of Toledo Law School. He has his law degree from University of Iowa and his Masters of Law from Harvard University. Professor Strang is chair of the Ohio Advisory Committee of the United States Commission on Civil Rights and

---

* Kurt Eggert is a Professor of Law at Chapman University Fowler School of Law, http://www.chapman.edu/our-faculty/kurt-eggert.
† Lee J. Strang is the John W. Stoepler Professor of Law & Values at the University of Toledo College of Law, http://www.utoledo.edu/law/faculty/fulltime/strang.html.
‡ Tom Campbell was the dean of the Fowler School of Law From 2011 to 2016 and is now the Boy and Dee Henley Distinguished Professor of Jurisprudence there, and Professor of Economics at the Argyros School of Business and Economics at Chapman, http://www.chapman.edu/our-faculty/thomas-j-campbell.

237
was a Visiting Fellow at the James Madison Program of Princeton University. Professor Strang is author of *Originalism’s Promise: A Natural Law Account of the American Constitution* and, also, the author of his own Constitutional Law textbook.

So, we will proceed today with fifteen minutes from Professor Eggert, followed by fifteen by Professor Strang, and then fifteen where the three of us will have a conversation, and then the last fifteen for questions and comments from the students. So let us introduce, with a warm welcome, our colleague, Professor Kurt Eggert.

PROFESSOR EGGERT:

Thank you. I appreciate all of you attending this debate. This is wonderful. I wrote the article we are discussing, *Originalism Isn’t What It Used to Be*, two years ago and never did I think that two years later I’d be addressing such a big and hopefully enthusiastic crowd about it. I have to tell you that I come to this issue from a different angle than many people who are debating originalism. Most people who do these debates are Constitutional Law professors and have an encyclopedic knowledge of Constitutional Law cases stretching back to antiquity, no, stretching back to the beginning of the Constitution. But I came to it because a couple of years ago we had a symposium on the nondelegation doctrine and *Chevron*—which are administrative law ideas—and I decided to write a piece on the nondelegation doctrine.

The more I researched the nondelegation doctrine, the more I became concerned, upset, outraged, worried, about the effect of originalism and how originalism was being used to turbocharge this doctrine which isn’t in the Constitution and which would give the Supreme Court great power over how Congress decides the administrative state should act. To some extent the nondelegation doctrine is a fight between the Supreme Court and Congress about how big and powerful the administrative state should be and how much regulation there should be of society. I want to frame this debate in those terms. I view originalism as a sort of philosophical buttress for the Supreme Court enforcing its policy preferences. And, in my view, the Supreme Court should not make policy. That’s the job of Congress, put in place by the Executive Branch. The Supreme Court should be enforcing laws and the Constitution, though doing so will naturally have policy implications. But, as so many people testifying in their confirmation hearings have said,
they should be calling the balls and strikes, not deciding what our country should be like.

I’d like to thank the Chapman Law Review for setting this debate up, it’s great. I’ve worked with them on their symposia. I’ve had a wonderful experience doing so. I’d like to thank Professor Campbell for moderating and Professor Strang for coming and providing me with this opportunity.

To frame the argument, let’s talk about a case that came out in early March. The Fifth Circuit decided the case United States v. Rahimi, involving whether people who are subject to a restraining order for domestic violence can, as part of that order, have their guns taken away which, as you can imagine, is a very important topic. The Fifth Circuit had previously said we permit this encroachment on the Second Amendment because it seems justified and workable, and therefore, it’s permissible. However, after the Supreme Court’s most recent Second Amendment case, N.Y. State Rifle & Pistol Ass’n v. Bruen, the Fifth Circuit said now the task is not to make this reasonableness determination but rather to see whether the restrictions put in place are sufficiently similar to historical restrictions on firearms and place a comparable burden on the right of armed self-defense. The court said you can only restrict guns in the same way restrictions were done when the Second Amendment was ratified. It has to be pretty similar. That case involved somebody who was involved in five different shootings and had allegedly assaulted and threatened his ex-girlfriend. She had received this restraining order by saying, essentially, “I’m in fear, he shouldn’t have guns.” So the court in her restraining order case took away his guns.

Now, if you look at the Second Amendment, it doesn’t say anything about restraining orders. The Court in Bruen said we have to look at what similar restrictions were in place back then, and whether the government, back in the time of the Second Amendment, took guns away from people it viewed as dangerous. The Fifth Circuit in Rahimi looked at such examples, parsed them out, and concluded that in each case, the historic examples were not similar enough to taking guns away from somebody subject to a domestic violence restraining order that we can abide by restraining order gun removal. So the Fifth Circuit held that the federal law banning the possession of firearms for specified

---

3 United States v. Rahimi, 61 F.4th 443 (5th Cir. 2023).
4 See United States v. McGinnis, 956 F.3d 747, 759 (5th Cir. 2020).
restraining orders is unconstitutional. In a stroke they took away important protections for victims of domestic violence. So that's the framework.

Let’s talk about originalism. Originalism is very difficult to define because there have been so many different forms of it. Here, the court was applying originalism as directed by the Supreme Court. But originally, originalism focused on the original intent of the Founders. What did the Framers think when they drafted the Constitution? However, that focus on original intent quickly fell by the wayside, or fairly quickly fell by the wayside, because people pointed out there’s no way to know what this collective body intended. Collective bodies can’t intend. We may know what individual Framers thought about specific issues, but we can’t say the Framers intended this just because Madison said it, or just because Hamilton said it. There was also a great article by H. Jefferson Powell in 1985 where he said not only that, but also that the evidence at the time indicates that the Framers didn’t want their original intent to be binding, but rather expected that the Constitution would be interpreted according to the plain meaning of the text, just like we do with everything else.6 And so, Powell said we shouldn’t have this idea that original intent governs, both because it is hard to figure out what that is, but also if the Framers didn’t want original intent to govern, if we want to honor their original intent, that means we should not do originalism. So, it was a pretty strong argument.

One can find quotations from the time of the founding that support both sides. Some indicate that the intentions of the Framers should have significance in interpreting the Constitution. Others indicate the opposite. I think it was Madison who said that the Framers of the Constitution should not be considered a great oracle for the interpretation of the Constitution.7

An important indication that the Framers of the Constitution did not want their intent to govern is how secretive they were about their discussions. You would think that if they were going to say, “our intent should be what people follow,” we would have had great records of the debates, which would be the best

---

7 Madison stated: “But, after all, whatever veneration might be entertained for the body of men who formed our constitution, the sense of that body could never be regarded as the oracular guide in the expounding the constitution.” James Madison, Jay’s Treaty (Apr. 6, 1796), in UNIV. PRESS OF VA., 16 THE PAPERS OF JAMES MADISON 290, 295–96 (David B. Mattern et al. eds., 1989).
expression of their intent. But instead, they met in secret, and their debates weren’t published until long after the Constitution was ratified. They didn’t put out to the public, “hey, ratify this only if you agree with our intent.” Instead, I think Madison’s notes, which are probably the best record of the Constitution, weren’t published until 1840, long after the original period. If they’d wanted their intent to govern, I think they naturally would have said “here is what we intend by this, here are the debates,” but they did not do that.

So, with these criticisms, the idea that it was the original intent of the Framers that governed was kind of blown up. So originalists scrambled around and said, “well it’s not the Framers we care about, it’s the Ratifiers, because it’s the ratification that really put the Constitution into place.” But there are bigger problems because there are a lot more Ratifiers, thus a much bigger collective body of which to figure out their intent. A lot of the records of ratification debates are either non-existent or really sloppy, so you can’t really ascertain what the Ratifiers intended, and, by and large, the Ratifiers were only asked one question of “do we ratify?” From that, you can’t conclude that they agreed to any intention about any particular statement in the Constitution. So then that way of originalism passed by and then we come to the next one.

Next emerged the third wave of originalism. This one was probably led off by Justice Scalia who said, in effect, “we should stop talking about original intent and start talking about original meaning.” So, the third wave is labeled original public meaning originalism, also known as the new originalism. I like the term “new originalism” because it is sort of like “jumbo shrimp.” The two words don’t go together. How can originalism keep changing if what it purports to be is the original understanding? The new original public meaning originalism was supposed to be more objective. We were not to depend on the subjective intent of the Framers or the Ratifiers, but rather on what the Constitution meant to the public. The claim is that the original public meaning is an objective standard. There are big problems with this, and this form of originalism is what I think a lot of originalists still use. Professor Strang has been studying this more so please correct me if I’m wrong in thinking that most originalists currently use some form of original public meaning. But the problem is how do you pin down how the public of the day understood the terms of the Constitution? Especially since many of those terms were written, I think, somewhat vaguely, as if to say we don’t know exactly what
this means, but they’ll figure it out as they go along in the future. Madison talked about how some terms of the Constitution will not have an exact meaning, either because of the difficulty of drafting or differences of opinion among drafters, and their more exact meaning will be settled and liquidated by practice as people decide on the meaning of vague terms, like cruel and unusual punishment. Do you mean cruel and unusual punishment at the time the Eighth Amendment was ratified? Or do you mean what is considered cruel and unusual punishment in the future when people are deciding that? There’s no way to use originalism to decide which approach to take. And so, people choose between them based on what they want to do.

The other problem with original public meaning is how do we figure out what people thought in 1781? As Scalia put it, that is an enormous challenge. It’s very difficult to put yourself back in the perspective of that distant time and forget everything you’ve learned in the modern day. And so, they say you look at old dictionaries, but dictionaries are terrible at interpreting something like the Constitution because they just give you the meaning of one word and, often, they give you multiple meanings of the same word, so you have to pick which one is most accurate. They rip the words out of context and they aren’t built for the context of the Constitution because the dictionaries at the time were written before the U.S. Constitution even existed.

One way some have tried to get around those problems was by saying, “well, let’s use the big data.” Something called corpus linguistics: taking a big corpus of a huge number of documents from the Founding Era and then using that to analyze what those old documents meant when they used the terms in question. But think about what we’re asking judges to do now. A problem with originalism is it calls on judges to do things for which they are untrained and have very little time to do. We’re asking them to be legal historians. We’re asking them to be linguistic historians – not only ask what the legal framework was then but also how people spoke and what they meant when they said things. Judges are really unable to do that in an effective way. Even Scalia, who was probably

---

8 Madison remarked:
It could not but happen, and was foreseen at the birth of the Constitution, that difficulties and differences of opinion might occasionally arise, in expounding terms & phrases necessarily used in such a Charter. . . and that it might require a regular course of practice to liquidate & settle the meaning of some of them.
Letter from James Madison to Spencer Roane (Sept. 2, 1819), in 8 THE WRITINGS OF JAMES MADISON 447, 450 (Gaillard Hunt ed., 1908).
the best legal historian on the Court, said judges aren’t given the
time or the research assistance to do that kind of historical
analysis.\textsuperscript{10} The Supreme Court is not really set up to do that. He said
it’s still better than any other system, but it is very problematic.

There are also a number of other forms of originalism that
have sprung up since then. This allows judges to kind of pick and
choose which form of originalism they want to use. Can they just
look at original intent if that helps them? Can they use original
public meaning if the original intent isn’t in their favor? They can
pick the form of originalism to use based on the policy outcome it
produces, which means that it’s not a constraint on their actions.
Originally, originalism was designed to restrain judges and force
judges to defer to Congress to set policy and only step in if it was
pretty clear that Congress had acted unconstitutionally. Originalism now, though, encourages courts to say “Hey, because
you violate our chosen idea of the original meaning, we can set
aside legislation kind of willy nilly.” Instead of judicial restraint,
what we have is judicial activism—though some conservatives call
it “judicial fortitude”—be brave enough to set aside laws that you
think violate the originalist view.

I have a couple minutes to get to nondelegation. The
nondelegation doctrine is based on the idea that Congress can’t
delegate to federal agencies the ability to make rules that bind the
public. It is an idea that has been bouncing around for a long while,
but has really been put back into play by Justice Gorsuch in a
recent dissent in the case \textit{Gundy v. United States}.\textsuperscript{11} He based
reinvigorating the nondelegation doctrine on originalist principles,
but the problem is if the Supreme Court can say no delegation,
that would mean that the Environmental Protection Agency
(“EPA”) or the Securities and Exchange Commission (“SEC”) or all
the people who are writing these rules that regulate private
conduct couldn’t write them anymore. Congress would have to
write the specific regulation that agencies have been creating.
However, Congress cannot write, in a timely fashion, regulations
to govern the environment because everything we know about it
changes so quickly. The purpose of this requirement is to make
sure there is a lot less regulation of private behavior.

What I argue is that the Framers hated the idea of judges
making policy. During the drafting of the Constitution, there was a
proposal that a Counsel of Revision, mostly made up of judges, would

\textsuperscript{10} See \textit{id. at} 860–61.

\textsuperscript{11} 139 S. Ct. 2116, 2131 (2019) (Gorsuch, J., dissenting).
veto legislation they didn’t think was good law, even if it wasn’t unconstitutional.12 The proposed Council of Revision was rejected because of the idea that courts should not be in the business of making policy. Courts are supposed to call balls and strikes, to overturn laws if they are unconstitutional, but not make policy and overturn laws just because they disagree with them. Now, the Supreme Court has become one of the central policymakers in the federal government. It wants to make policy on what regulations the EPA can make. It wants to make a great amount of policy based on its originalist conceptions and doing so causes the Court to act in a manner that the Framers in the founding era never intended. They never wanted people to turn to the Supreme Court for major policy decisions, to say, “what is the rule on guns? Well, it’s up to the Supreme Court. What’s the rule on global warming? Well, it really depends on what the Supreme Court says.” They would be appalled by the idea that the Supreme Court has become one of the primary policy making bodies in Washington, but that’s where we are.

I think my time is up, so thank you.

PROFESSOR CAMPBELL:

Happy to welcome Professor Lee Strang. Thank you, Professor Strang.

PROFESSOR STRANG:

Thank you so much. Good afternoon. Great to be with you here at Chapman. It’s my first time at the law school. Beautiful law school, beautiful campus. Thank you, Professor Eggert, for inviting me to come and debate. He and I first met virtually a couple of years ago when he published the article that was published with the Chapman Law Review. I read it. He and I corresponded. I gave him some thoughts. Good article, I thought. I learned from it, so I appreciate it. And then he reached out and said, “let’s do an exchange like this.” So, I really am looking forward to learning from our exchange.

Our debate today is—does originalism work? And there are a number of ways that question could be interpreted. The question could mean, does originalism work based on its own terms? That is, originalism tells us a story. Here’s how you do it. Professor Eggert, I thought you did a great job talking about three to four different

instantiations of how to do originalism. Originalism tells us a story of how we do it, and is originalism actually able to live up to its promise about how to do it?

There's a second way to interpret the question: even if originalism lives up to its goals—in other words, original intent, original meaning, whatever—is that the right way for us to interpret our Constitution? I'm going to focus in my initial remarks on the first question, and then in my conversation point, I want to respond to Professor Eggert’s direct criticisms as well. I'm looking forward to your questions and comments. There are a lot of things that I can't say in fifteen minutes!

My argument this afternoon, in my initial remark, is that originalism, as it promises, is an effective means to faithfully follow and implement our written Constitution. In some ways, when you think about originalism, these are the best of times to be an originalist, at least since the New Deal. When you look at nominee Kagan’s remarks during her 2010 confirmation hearings, my living constitutionalist friends were scandalized when she said, “we’re all originalists now.” And just this past year, nominee Brown-Jackson stated during her hearing, “I am focusing on the original public meaning because I'm constrained to interpret the text.” And she elaborated: “I do not believe that there is a living Constitution... in the sense that it’s changing and it’s infused with my own policy perspective or the policy perspectives of the day. The Supreme Court has made clear that when you’re interpreting the Constitution, you're looking at the text at the time of the founding.”

I think that’s strong evidence that originalism is really becoming a powerful force, not just on the Court, but elsewhere.

My remarks are meant to provide reasons why it is that people like Justice Kagan and Justice Brown-Jackson, who you might think are not inclined towards originalism, at least as it’s currently or conventionally understood, still describe themselves as originalists. So, first, to know how originalism works by its own standards, we need to know what originalism is.

Originalism in its modern scholarly form came to the forefront in the 1970s. Judge Bork was an original intent originalist and then segued over to original meaning originalism, as Professor Eggert talked about. Since that time, originalism has grown in sophistication and influence. I think a recent powerful piece of evidence was Adam Liptak, *By Turns Cautious and Confident, Judge Jackson Takes the Stage*, N.Y. Times (Mar. 22, 2022), http://www.nytimes.com/2022/03/22/us/ketanji-brown-jackson-judicial-philosophy.html.
evidence is the Dobbs case. When you compare the Roe Court, where you had two dissenters, to the Dobbs Court, where you had six to three—six originalist justices overruling or limiting—you see the influence of originalism on the Supreme Court.

One other thing I would suggest is that the Supreme Court oral arguments are available publicly on the Oyez website. Go back and listen to the Roe oral argument and then compare it to the Dobbs oral argument. And I think not only do you see a change of outcome; what you see is a change of focus of what the judges and the advocates think it is they're focusing on. Compare the Roe Court’s way of thinking about what the role of a court is and then compare it to that of the Dobbs Court, and I think it shows you the influence that originalism has had in focusing legal argument on text, structure, history, and precedent.

Second, how do you do originalism? Originalism in principle is the idea that the public meaning of the text, when it was ratified, is the Constitution's authoritative meaning. So, let’s say that we were trying to find out the original meaning of the word religion in the First Amendment. We’d see that the word appears twice, first in the Religious Test Clause in the original Constitution and then also in the First Amendment. And so, we know that the text says the word “religion.” And we also know that the Framers and Ratifiers were speaking a language that looks a lot like our language. But you also know, especially if you’re an English major, that there’s a phenomenon of linguistic discontinuity: that in natural languages that people are speaking—living languages—the language’s meanings change over time, haphazardly and unexpectedly. And so, we can’t just rely on the text and today’s conventional meaning.

Next, we look at the structure of the Constitution. We see that in every instance where the word “religion” appears, it’s as a limit on the federal and then later state governments. That gives us more information. It tells us that religion is something that the government is interested being involved with. And it tells us that the American people in two instances said, “no government, stay hands-off of this phenomenon called a religion.” But it doesn’t give us enough information to answer a lot of questions.

Third step: you look at the framing ratification debates. Now, the ideal would have been if James Madison, when he introduced in 1789 what became the Bill of Rights, had said “by the word religion, I mean...” and then went on to define it. But he didn’t do that, and we shouldn’t expect him to do that because he was speaking English in the same way that you and I speak English.
We don’t define the terms that we’re using. Instead, he relied on
the language conventions in use at the time. And so, what the
originalist will do is look for every time where the word religion
was used in the first session of the first Congress and when the
First Amendment was introduced in the state ratification
conventions, and see what was the conventional meaning for the
word religion at that time.

Next step, expand the data set, because what we’re looking for
isn’t simply what James Madison thought the word religion
meant. We’re finding out what was the public meaning of the word
religion in 1791. And in principle, you can find evidence for that
public meaning anywhere where conventional English was used in
the United States at that time period. It can be in speeches,
sermons, state statutes; it can be in private letters between people
who are facile with the use of the word religion.

And then lastly, you look at the cultural, philosophical, and
religious context and you ask: in a political community like this,
with its commitments and its understanding, its concepts, what
would those folks understand the word religion to mean? And the
reason why I picked the word religion was because, in a series of
articles, I went through an originalist inquiry for the word
“religion.” And my conclusion, for what it’s worth (we can talk
about it later), was that religion is a belief system with belief in a
deity, with duties in this life—thou shalt and thou shalt nots—and
a future state of rewards and punishments. So that’s what I’ve
argued the original meaning of the word religion was.

Last comment—corpus linguistics. Professor Eggert had
identified corpus linguistics, and it’s a tool that originalists
identified maybe five or so years ago as an additional way to help
provide information for and checking of originalist scholarship and
research. And so, after I did research on the original meaning of
religion using the contentional techniques, I went back and did a
corpus linguistics search to try and uncover additional evidence
about the word religion. And what you do is you utilize massive
bodies of electronically searchable documents. They’re called
corpora, and then you search them for every instance where the
word religion is used. And then you can use the tools of corpus
linguistics. It has its own terminology and tools including things
like “co-location”: what word appears most often with the word
religion within five words? And then what you are able to find is
that the language practice at the time, what did people, when they
used the word “religion,” think it was similar to or synonymous
with? So, corpus linguistics is another tool that originalists are using to make sure that they’re more accurate.

Up to this point, you might say, “okay, I think I know how to do originalism,” but our legal system, as law students know from when you take your first-year courses, has a lot of precedent. There’s a lot of stare decisis. And I haven’t said anything about stare decisis. In fact, what I’ve said seems inconsistent with stare decisis. Because what I have said is that you identify the original meaning and you follow it, and there’s not a word about stare decisis. Indeed, many critics have plausibly argued that originalism’s commitment to following the Constitution’s fixed original meaning means that adopting originalism would lead to the overruling and destabilization of broad areas of American law. Confirming critics’ worse fears, Justice Thomas recently argued, in a concurrence in the *Gamble* case in 2019, that originalism doesn’t quite have no space for stare decisis, but it’s a really small space. So, he’s basically saying, “we don’t do stare decisis around here.” But I don’t think that’s right. My view, as I argue in my book *Originalism’s Promise*,¹⁴ is that federal judges create and, in turn, are required to follow constitutional precedent because the Constitution itself commands that they do so. The original meaning of judicial power in Article III, the power federal judges utilize, requires them to follow precedent. So, the very first questions that a federal judge asks when deciding a case is, “is there precedent on point?” And then there’s a little bit of nuance if the precedent is an originalist precedent—that the precedent fully identified and articulated and applied the original meaning versus a non-originalist precedent. But even non-originalist precedent in some instances, in my view, will be followed.

So, at this point, you might say, “okay, I understand how originalism might operate. I understand how it incorporates stare decisis. Stare decisis plays a big role in originalism.” But, as the title of our discussion asks: does originalism work well meeting its goal? Over the past twenty years, originalists have articulated a sophisticated and nuanced approach to constitutional interpretation, one that simultaneously gives pride of place to identifying and following the original meaning, which is what I’ve just argued, while also recognizing that the original meaning may not always clearly answer a question. So, originalists have identified the Constitution’s own mechanisms

to implement the Constitution’s original meaning, even when it might not be clear initially.

First, it’s important to note that my view is that there is significant consensus on most of the original meaning of most of the important provisions in the Constitution. These include things like Article I, Section 7 – how Congress creates law. They include the Interstate Commerce Clause, the Necessary and Proper Clause, the Establishment Clause, the Privileges or Immunities Clause, and the Cruel and Unusual Punishments Clause. Professor Eggert had mentioned “cruel” from the Eighth Amendment. I think there’s a deep and broad consensus on the original meaning of all those provisions and others.\footnote{To be clear, my claim is that there is a broad consensus among scholars about the original meaning of these (and other provisions), not that all of the meaning of those provisions has been liquidated.} I picked those because you covered all, or almost all of those in your Constitutional Law classes, and I think that there is a deep consensus on them. Professor Eggert mentions the nondelegation doctrine in his remarks, and his article—a good article—is about the nondelegation doctrine. I think there’s a consensus on that too, and I’d be interested to see Professor Eggert’s response. The consensus is not that there’s no nondelegation doctrine, and not that the nondelegation doctrine is something where administrative agencies can’t do anything. I think it’s something in the middle, and I think there’s reasonable debate about what that something in the middle is. The consensus is that Congress can delegate some power, but not unlimited power. Now, maybe in my remarks later I can say a little bit more about where I think that line is at.

This doesn’t mean there’s unanimity. My claim isn’t that everybody thinks the Establishment Clause means X or Y, but I don’t think unanimity is the standard for any human practice, because we humans have a penchant for disagreeing. So instead, I think the standard is—the relevant metric is—a consensus among scholars.

Second, originalism has identified four methods to identify further consensus. So, in other words, even if there’s not a consensus right now, here are four ways to identify further consensus. First is the method of triangulation. The method of triangulation has three distinct ways of identifying the original meaning. And the key is, if all three ways point towards the same original meaning, you’ve got a lot of confidence that you’ve arrived
at the correct original meaning. If they point in different directions, you don’t have much confidence at all. You need to go back to the drawing board and try again and in different ways.

The methods of triangulation include historical immersion, where one immerses oneself in the conceptual world of the Framers and Ratifiers. Second, studying the record, where one reads the primary sources of drafting and ratifying constitutional text. And third is corpus linguistics, which is something that both Professor Eggert and I talked about.

The second tool to build consensus is called the scholarly division of labor. Professor Eggert, I think, appropriately criticized the Supreme Court justices who talk about themselves using originalism to identify limits and meanings of the Constitution because, as he said, they don’t have time and they don’t have the experience. But through the scholarly division of labor, scholars research and debate and come to consensuses about what the original meaning is, and then judges use that consensus. Justice Thomas does this. So, for those of you who have taken Property and studied the Takings Clause, you know that Justice Thomas, in his dissent in *Kelo v. New London, Connecticut*, relied on originalist scholars to identify the original meaning of the Public Use Clause.

Third tool: scholars have developed consensuses on many areas of constitutional meaning. The consensus has three points, actually. First, there’s consensus about what we agree on, so you can have a high degree of confidence that the Constitution includes that as the original meaning. Second, there’s consensus about disagreement. In other words, we know we don’t agree on a proposition. And then, third, there are areas about which there’s actually disagreement about what scholars (dis)agree on. That’s a really deep area of disagreement. And the detailed example I would give of each of these three aspects would be the Privileges or Immunities Clause, if we had more time.

Fourth are closure rules. An important example is called the best-available-legal-evidence rule, and here’s how it works. If you’re a Supreme Court justice presiding over a case, you have two parties before you. One party says the Constitution means X, and the other party says it means Y. And the justice has to make a decision—that’s what Article III requires—so the justice should rule for whichever party has presented the best available legal evidence. Maybe not the best evidence in all possible worlds, maybe it’s not the best in the overall scheme of things, but the best of the two arguments being presented to the justice in that case, and the justice makes a decision based on that. Is that something
judges can do? The answer is clearly yes, because that’s what they do on a regular basis.

Third, originalism identified institutional mechanisms to resolve remaining underdeterminacy. The two institutional mechanisms are stare decisis and constitutional construction. Because I am nearly out of time, I will say one thing only about constitutional construction. Constitutional construction is the idea that there are times when the Constitution’s original meaning runs out. Originalists agree with what Professor Eggert is saying. And in that situation, what originalists have said, is that the relevant interpreter needs to “construct” constitutional meaning.

I actually think the nondelegation doctrine may be an example of construction. I think that there are two propositions that are clear: (1) Congress can delegate some powers to agencies; and (2) Congress can’t delegate all its power to agencies. And the relevant rule, which we can talk more about in our conversation, is somewhere in the middle, and that the precise contours of how that relevant rule is applied might be an area of construction: where the original meaning is underdetermined. And so, the relevant interpreter, the Court or Congress, is the one that constructs constitutional doctrine.

In conclusion, I made two moves this afternoon in my initial remarks. First, I described originalism. Second, I showed that originalism is faithful to its commitment to the original meaning and sophisticated in its approach to implementing that original meaning. Thank you very much.

PROFESSOR CAMPBELL:

So this is the interrogatory part. We are going to give the first comment to Professor Eggert to raise any subject he wanted to from Professor Strang’s remarks, and then we will give the reciprocal privilege to Professor Strang. So, Professor Eggert.

PROFESSOR EGGERT:

Thank you. I appreciate your comments and I think that you have referred to one of the most important, but little discussed, issues that originalism faces, which is: If the Court is facing a question where even the Court thinks it’s kind of muddy what the original public meaning was, when should the Court act? So, for example, if the Court thinks it’s really unclear what the original public meaning of a term of the Constitution was, but then there’s a fifty-one percent chance that the public meaning means “X,” and hence, Congress’ act is unconstitutional, should the Court
overturn that act? I think that's fundamentally wrong, that the Court should not be overruling—the Court should initially defer to Congress to say “this is a tough area, you're Congress, you're the one who's setting policy. If we're not sure that what you're doing is unconstitutional, we should just let it go.” The Court should just say “okay, that’s your decision. We can't say for sure that you’re wrong.” But what we see from the Supreme Court now is application of the best available legal evidence rule. If the Court concludes that the best available legal evidence, which may not be that great, indicates that maybe the act is probably unconstitutional—the Court can overturn legislation it disagrees with for policy reasons. That allows the Court to find legislation unconstitutional based on flimsy evidence of original public meaning, in order to make policy decisions that really usurp Congress's power to make policy through its legislative power.

PROFESSOR STRANG:

Thank you Professor Eggert for that question. So, Professor Eggert described, as I understand it, a situation where an action by Congress is being challenged as being beyond Congress' power or violating some constitutional right. And so, under the best-available-legal-evidence standard, a court could have a low degree of confidence of Congress not having the power or the right being infringed, and still strike down Congress' action.

The way I would evaluate this situation, as an originalist, would be: I would find out what judges are authorized to do by Article III, because if, in fact, Article III authorizes them to do as Professor Eggert had said—which is to defer to Congress in areas of, let's say, low certainty—then I don't think it's a question of ethics, I think it's a question of law, and they should do exactly what Professor Eggert had said. From my perspective, then, it's simply a question of what does our law in fact require. And I made the argument in my initial remarks that our Constitution in fact requires deference to Congress in areas of stubborn constitutional underdeterminacy.

Just to be clear: I think it remains an open question whether to and what degree federal judges should defer to Congress. I have good friends who are scholars in this area who say that the original meaning of “judicial power” requires judges to defer to congressional judgments, unless there is a clear error by Congress; so in other words, their argument is a historical argument about the original meaning of Article III. I haven't been persuaded yet, so I think that's an area where there's a debate, but they are thoughtful scholars,
and they might ultimately be right on that point, and if they are, then that would support Professor Eggert’s position.

There’s one more thing I would say about Professor Eggert’s concern. His concern is Congress won’t be able to act. But actually, it doesn’t always turn out that way because there can be situations where, for example, it can be a state versus an individual, or it can be Congress versus a state, or it can be two states versus each other. And so, there are lots of instances where there is constitutional underdeterminacy, where there is a low-degree of confidence about who is ultimately right, and Congress wouldn’t be a loser one way or the other. So I actually don’t think that the concern you have applies to all, or perhaps even most, situations.

PROFESSOR CAMPBELL:

Professor Strang, do you have a point you want to bring up relative to Professor Eggert?

PROFESSOR STRANG:

Yes. The burden of Professor Eggert’s remarks was that originalism was this malleable, evolving theory from which judges can pick and choose. And I think he graciously didn’t say they intentionally do it. I don’t think he said that people are intentionally misinterpreting the law to achieve their policy goals, which I think is probably true. When I look at the justices, I view them as mostly acting in good faith, even if I disagree with them. But what ends up happening, according to Professor Eggert, is that the originalist judges choose things like the nondelegation doctrine and use that as a way to limit Congress in an inappropriate way. I don’t think that’s true.

I believe the consensus of the nondelegation doctrine isn’t that there’s no doctrine—Professor Eggert had called the doctrine a myth, so I think that might be his position. The consensus isn’t that the nondelegation doctrine means that Congress can transfer all of its power.

The consensus is in the middle. And I think that middle is actually represented by a case many of you have read. Those of you who have had administrative law, you probably read Wayman v. Southard, an 1825 opinion by Chief Justice John Marshall. Marshall says basically two things. Congress can transfer power to the executive branch or to the judiciary, and these branches can decide unimportant issues. Now, what is unimportant I think may be part of the construction zone. For example, Congress can give to another branch the power to fill in the details of an important
issue that Congress has decided. I think a position like that is actually where most originalist scholars are at.

This seems to me to be a reasonable position because, under Professor Eggert’s view, which is that the nondelegation doctrine is a myth, nothing would prevent Congress from passing a statute, we’ll call it the Goodness and Peace statute. The statute directs the new Goodness and Peace administrative agency run by former-Governor Newsom and now secretary Newsom, to create regulations for the goodness and peace of Americans. And then Congress decides to go into recess *sine die*. And so, since there’s no nondelegation doctrine, in your view, it seems like they can do that.

**PROFESSOR CAMPBELL:**

Professor Eggert gets to respond and then we’ll open it up a bit.

**PROFESSOR EGGERT:**

I think that’s a great hypothetical, and I’ve had several people propose it to me and here’s my answer to that. You cannot find the nondelegation doctrine in the Constitution, so if you’re an original public meaning person there are no words to attach it to and so the nondelegation doctrine should not be enforced. However, the Constitution does vest legislative power in Congress. So, that is an express provision that we can interpret. The hypothetical that you mentioned, if it were to happen, I think violates the Vesting Clause because Congress can delegate powers to federal agencies, but it can’t vest its legislation powers in federal agencies. And, if the Congress passed a law as you described, I think that would be a re-vesting because Congress would not have the power to ride herd over what the agency did or to pull back or to change laws. It would just have transferred its legislative power away which I think violates the Constitution’s vesting clauses by vesting legislative power somewhere else rather than in Congress.

**PROFESSOR CAMPBELL:**

Go ahead Professor Strang.

**PROFESSOR STRANG:**

I’ll just change the hypothetical slightly. So, Congress didn’t go into recess forever. It went into pro forma recess for C-SPAN viewers like me, once every six months, and then the Speaker of the House says to the empty chamber, “Is Secretary Newsom doing a good job with peace and goodness?” and the answer is “yes.”
PROFESSOR CAMPBELL:

I think your hypothetical works even if you don’t have the *sine die*, so you’d probably want to take that out. I’m going to interrupt you, to be aggressive because it is within my nature. I have an aggressive question to both. First of all, to Professor Strang and then I’ll give you your question too, Professor Eggert. Professor Strang, tell me one example where an originalist, using the originalist approach, comes to a conclusion that she or he would not have on policy grounds. In other words, I’d love to know if this is a null set, does it actually constrain, or is it just a makeweight? I have actually one to give if you don’t. Professor Eggert, so what are the constraints? Admittedly, originalism isn’t perfect. None of these sources are perfect. Some statutory constructionists debating what the committee said and what was said on the floor, it’s not perfect but it’s something. And, as Professor Strang put with his hypothetical, you have these mores, the word that Justice John Paul Stevens used in the case striking down the use of exclusionary race by colleges. Somebody must determine what the mores are. Are you comfortable with the Supreme Court deciding what the mores of our time happen to be? So, first question to you.

PROFESSOR STRANG:

So, I think the evidence of my answer to Professor Campbell’s question is ubiquitous. I’ll just speak for myself. There are a number of aspects of the Constitution’s original meaning that I don’t think are wise, and some of them not even just. I’ll pick one example. I think the free exercise of religion should be robustly protected, but I don’t think the original meaning of the Free Exercise Clause protects it very robustly. I think *Employment Division v. Smith* was probably right. The author of *Smith* is another example. Famously, the author of *Smith* was Justice Scalia, an originalist, and he articulated a decision that was relatively unprotective of religious liberty.

And even on the point that we were talking about, nondelegation, Justice Scalia authored what I think is probably still the most recent important Supreme Court decision on the nondelegation doctrine, *Whitman v. American Trucking Association*. In his majority opinion, he effectively said that the nondelegation doctrine, outside the context of the National Industrial Recovery Act—those are the *Panama Refining* and *Schechter Poultry* cases from the New Deal—and the *Mistretta v. United States* delegation of pure legislative power to the Sentencing
Commission, there is not an enforceable nondelegation doctrine, and I am very confident he does not think that is a wise decision.

PROFESSOR CAMPBELL:

I was going to give you Justice Stevens’ concurrence in the flag-burning opinion.

PROFESSOR STRANG:

Or Justice Scalia’s!

PROFESSOR CAMPBELL:

Or Justice Scalia’s too. And to you Professor Eggert, so how do you keep the mores from just being decided by the majority of the nine?

PROFESSOR EGGERT:

You posed two questions to me. The first question was: If we don’t do originalism what do we do? And I think the answer to that is that we do care about what the original intent seems to be, even though we recognize the issues with determining it. We do care about the original public meaning, but we don’t say it is dispositive, that it is binding. We say that that the original intent and public meaning create a great jumping-off point that we always have to be cognizant of, but we also have to look at how society has changed. We have to look at how the law has developed, how other laws have developed. We have to try to put it in the context of today to recognize the ideals and principles of yesterday. But we have to contextualize it, and I think that’s where originalism often breaks down. In the Second Amendment case involving firearm restrictions in restraining orders, they didn’t address the fact that we didn’t have domestic violence restraining orders back then that removed guns, in part because domestic violence wasn’t really frowned on that much. Domestic violence was not illegal everywhere and restraining orders for domestic violence didn’t exist as we know them today, and so, of course they didn’t have rules limiting firearms for domestic violence restraining orders back then. We shouldn’t say that if they didn’t have rules like that back then we can’t have rules like that now because that is too much a constraint on the popular will which I think does support such rules. As to the mores of society, I think that’s something that Congress should have the first crack at because members of Congress are representatives of the people; if they get off track from what the people want, they’ll lose an election. Supreme Court justices can sit for thirty or forty years, they have little worry about
popular opinion and so I think are much less reflective of the popular mores of the day, and that’s a real problem.

PROFESSOR STRANG:

I’ll just say one thing. Professor Eggert, in his initial remarks and just now had said one of the challenges with originalism, one of the problems with originalism, is that it’s a cover for the justices’ policy decisions. So, remember what someday-justice, but right now Professor Eggert had said.

PROFESSOR EGGERT:

Not going to happen...

PROFESSOR STRANG:

He said: we care about the original meaning but it’s not dispositive, it’s a jumping off point. So, think about what we’ve been debating. I’ve been saying that, within limits, one can identify and follow the Constitution’s original meaning. Professor Eggert has given us a standard where the original meaning, it seems like it matters, but we don’t know how much, we don’t know to what extent, and we don’t know based on what reasons one doesn’t have to follow it. It seems like what characterizes—in other words the characteristic that is central to his version of living constitutionalism—is judges making their own decisions. Whereas originalism is characterized by following the original meaning.

PROFESSOR CAMPBELL:

The description of Congress being subject to the election every two years, and hence a natural check, speaks to where Congress is acting against the popular will. But Congress often acts completely consistent with the popular will, just in some ways which are abhorrent to some constitutional principles, such as segregation for many, many years, quite consistent with the popular will in those states that had it. And so, I put forward to you a third possibility, and that is amending the Constitution. We have not discussed that today, but it seems to me a very important part of the separation of powers doctrine. So, rather than say the Congress will eventually correct itself, because it won’t, we could amend the Constitution. No one ever gets re-elected by saying, “I’m standing for criminal rights. I am going to work hard to get more people out of jail earlier.” You don’t win elections that way. And similarly, I mistrust a justice who has no guardrails, and obviously, Professor Eggert, you’re not suggesting this as well.
But suppose you say the Constitution can be amended and that is the ultimate of both because it is the popular will, admittedly a high bar. For example, the Second Amendment says “[a] well regulated Militia, being necessary to the security of a free State, the right of the people to keep and bear arms, shall not be infringed.”\textsuperscript{16} That means you’ve got the right to get a bazooka because the Framers intended to get a bazooka to take out the national government if they became onerous. But we don’t want that today, so let’s amend the Constitution. It seems to me that the amendment process is ignored as we divide up, is this going to be Congress or is it going to be the courts? And lastly, for those who say it never happens, when I was in college, I couldn’t vote. The Supreme Court said that the federal law allowing eighteen-year-olds to vote was unconstitutional. That decision was in December of 1970. In July 1971, fewer than seven months later, the Constitution was amended. It can be done if there’s a consensus.

PROFESSOR EGGERT:
I want to get to questions.

PROFESSOR CAMPBELL:
Very good, no rebuttal.

PROFESSOR EGGERT:
You can’t rebut that.

PROFESSOR CAMPBELL:
Alright so I think I should choose questions from this side and then we’ll go over there. Please go ahead sir.

QUESTION ONE:
So, I would like to know from all three of you, do you think that the ability to amend the Constitution is still a practical reality? And as a follow-up to one little remark you said about the impossibility of getting elected on the platform of criminal rights, I am pretty comfortable that that’s false. I’m from Philadelphia, and that is one of the places where it is very much possible to get elected on a reform platform or an abolition of police imprisonment platform.

PROFESSOR CAMPBELL:
I’ll give this to my colleagues because I’ve already spoken on this.

\textsuperscript{16} U.S. Const. amend. II.
PROFESSOR EGGERT:

I’d like to jump in. I think amendment is possible, but what we’re talking about is interpretation of the existing words, and how existing words are interpreted is almost impossible to change by amendment. For example, with cruel and unusual punishment, there’s a great disagreement on what those words mean. What we consider cruel and unusual today or what was cruel and unusual back then? How do you amend the Constitution to change the meaning of the words if you want to keep the words as is?

PROFESSOR STRANG:

I think it is very hard to amend the Constitution on anything about which Americans reasonably disagree, which in today’s world, because of polarization, is virtually everything of importance. I actually think that my friend Professor Eggert’s view of interpretation is partially to blame for that. I don’t think it’s a coincidence that the New Deal is when Americans stopped adopting important amendments. We, of course, changed voting rights from twenty-one to eighteen, and I think that was supported by reasons, but I don’t think it’s of the same importance as how do we elect senators? Should we have non-discrimination in voting? Should we have a progressive income tax? Those are what I would think are fundamental and transformational. We don’t do that anymore.

And here’s a way that you can test it yourself. When it comes to presidential elections, what do you say to your fellow citizens? You say, I bet, “vote for this candidate or against this candidate,” in part because of the justices that that person will nominate. When’s the last time you said to a fellow citizen, “vote for or against a particular constitutional amendment?” I think the answer for most of us is zero, if at all? And so, that suggests to me that, because we all know what the “game” is, we’ll just get our justices appointed to the Supreme Court where they will interpret the Constitution the way we want to.

PROFESSOR CAMPBELL:

We amend our Constitution in California every two weeks. Question from this side? Yes sir.

QUESTION TWO:

Professor Strang, you stated that Professor Eggert’s viewpoint is characterized by justices’ discretion. It seems like, in one of his law review articles, Justice Scalia suggests that we would abandon originalism if there was a true bitter outcome.
And so, I think he was talking about cruel and unusual punishment there, and so what makes that perspective different from Professor Eggert’s?

PROFESSOR STRANG:

Justice Scalia was one of the leading lights for originalism for a long period of time. That doesn’t mean that I agree with everything that Justice Scalia has said. He’s a human being, so he makes mistakes. Justice Scalia later came to regret the statement that you are identifying, because I think he would agree with the position that I’m going to take now, which is that: when he swore to uphold the Constitution, he swore to uphold its original meaning. And if one is put in a position where one’s oath requires one to take actions that, in fact, are deeply unethical, then one has, I think, one or two options: resign from the office that one is in or, if possible, recuse oneself in some way from taking that action (if allowed by the law). That is a summary of the position that I take.

And I think that position is widely shared among people. In other words, if you’re a judge, you’re bound to run into situations that are ethically tenuous. If you’re against the death penalty, for example, and you think that the death penalty is constitutional, then you are going to come into ethical conflict. And so, I think the right originalist approach is to be up front about the potential for conflict.

I think what living constitutionalism invites people to do is what the Supreme Court did in the 1970s, which is to “creatively” interpret the Eighth Amendment to say the death penalty has always been unconstitutional, even though we never knew it.

PROFESSOR CAMPBELL:

From this side, please.

QUESTION THREE:

So, you talked about corpus linguistics in your talks. Every year we have breakthroughs in sentiment and intent analysis through natural language processing. One of those breakthroughs has been 3D vector analysis where we sort words based on a 3D vector space where each axis assigns to sentiment, or intent, or something like that. To what extent do legal scholars or judges have a duty to use these tools? And similarly, if they do have a duty, either ethical or legal, what duty do they have to use this for normal statutory interpretation, using a corpus from the sixties to interpret law written in the sixties.
PROFESSOR EGGERT:

I think that’s a great question. I’m interested to hear how Professor Strang responds. But I think this shows the impossibility of this task because very few judges would understand the process of using corpus linguistics that you’re talking about. And if you’re a judge, how can you say, “Well I think that this is constitutional based on my legal reasoning, but this program spit out something that indicates otherwise, I don’t really understand it, but I still have to go with it.” That is enormously problematic.

PROFESSOR STRANG:

Corpus linguistics is a subset of something that humans have experienced, and Americans in particular, for as long as we’ve been around: which is technological change. And normally, we say there is a new tool, a technological change that’s allowed us to do something we’ve already being doing, but do it better. We normally say it might be more complicated, it might be more challenging, but we should embrace it because it allows us to do what we’re already doing in a better way. So, my short answer would be, yes, if the addition to corpus linguistics allows us to better understand whatever it is we’re aiming for, let’s use it. I don’t think judges typically should be doing it unless they have special training, which is unusual. I think what they need to rely on is a community of scholars who, through debate and discussion, come to a consensus on what they agree on, a consensus on what they disagree on, and then there is remaining dissensus.

PROFESSOR CAMPBELL:

I wish to note that we have a scholar at our law school named James Phillips who is nationally recognized as a leader in the field of corpus linguistics.17 So perhaps on another occasion, we’ll hear from him. But it’s really true, if you go across the country, you won’t find a better cited or respected scholar in just this field than my colleague James Phillips. Now we’re at the point now for concluding comments so I’m going to allow the first comment to come from Professor Eggert and then to our gracious visitor to have the last word.

---

17 Associate Professor of Law, Dale E. Fowler School of Law, Chapman University, http://www.chapman.edu/our-faculty/james-phillips.
PROFESSOR EGGERT:

My central thesis is that originalism has become a method by which the Supreme Court can justify overturning legislation it disagrees with for policy reasons. And the Court seems increasingly willing to do just that and seize policy-making power from Congress. The original originalists would have been horrified generally at the idea that the Court should set policy, because originalism was originally designed to enforce judicial restraint. Now, instead of judicial restraint, originalism is being used by the Court to make policy decisions, and it makes me worried about what will happen in the future.

PROFESSOR STRANG:

We should try and interpret our Constitution the way that, by our own lights, we think is supported by sound reasons. If originalism is the way to do that, then we should do originalism, and if not, then not. As I've written in my book, Originalism’s Promise, I think originalism is supported by sound reasons, and so to the extent that we can identify a consensus on the original meaning, I think judges should follow it. And to the extent they are not making policy, what they’re doing is following the Constitution in the way that we currently think is best supported by reasons.

PROFESSOR CAMPBELL:

We are deeply grateful to the Chapman Law Review for allowing us to come together and creating this opportunity, specifically to discuss the issue of originalism but other subjects, no doubt, in the future as well. We’re so grateful to the author of the article who gave us the opportunity, Professor Eggert, and to Professor Strang with your insight and experience that brought such light to this subject. Will you all join me in thanking them?
H-1B Visa Lottery: Random Selection Process Prevents Applicants from Achieving the American Dream

Alexandra Amos

“Once I thought to write a history of the immigrants in America. Then I discovered that the immigrants were American history.”

– Oscar Handlin

INTRODUCTION .................................................................264

I. OVERVIEW OF THE H-1B VISA PROGRAM ....................267
   A. H Category Visa Before 1990 .................................267
   B. Inception of the H-1B Category ...............................268
   C. H-1B Petition and Selection Process .......................270

II. INTENT OF THE H-1B PROGRAM ..................................272

III. ISSUES WITH THE H-1B PROGRAM ...........................274
   A. Issues with the H-1B Yearly Cap ............................274
   B. Misuse of the H-1B Program .................................276

IV. BENEFITS OF THE H-1B PROGRAM ............................278
   A. Innovation Benefits .........................................278
   B. Economic Benefits .........................................279

V. MERIT-BASED SOLUTION ...........................................281
   A. Wage-Based Selection ......................................281
   B. Points-Based Selection System ............................283
   C. Foreign Points-Based Immigration Models .............285
   D. Employer-Driven Models ..................................287
   E. Measuring Merit Depending on the Country’s Immigration System ..................................287

* J.D. Candidate, Expected May 2023, Chapman University Dale E. Fowler School of Law. B.A. Lawrence Technological University, 2020. Special thanks to the inspiration of this article, Santiago Amigo. America is lucky to have you.

VI. BENEFITS OF MERIT-BASED PROGRAMS

A. Merit-Based Selection Process Bolsters Intent of the H-1B Program

B. H-1B's Structure is an Ideal Candidate for Merit-Based Selection

C. Measuring Merit

D. Problems with Merit-Based Systems

E. H-1B Program Answers the Problems of Merit in its Structure

CONCLUSION

INTRODUCTION

The American Dream: “a happy way of living that is thought of by many Americans as something that can be achieved by anyone in the United States especially by working hard and becoming successful.”

Hard work. Perseverance. Opportunity. These are words often associated with the concept that anyone who works hard, is driven, and that if one plays by the rules, they can achieve success and stability in America. The original concept of the “American Dream” was founded in ideals of equality, justice, and democracy. However, many hardworking, driven immigrants would probably disagree with the idea that a path to success in America can be paved simply through hard work and success. Across the United States, people mistakenly believe that immigrating to the country “the right way” is an easy or even attainable process. However, there are very few ways to immigrate to the United States legally. The United States offers green cards to four basic categories of individuals: (1) relatives of legal immigrants; (2) a limited number of skilled workers

---


6 Id.
How The H-1B Visa Lottery Prevents the American Dream

sponsored by U.S. employers; 7 (3) refugees and asylum seekers; and (4) applicants from countries that send few immigrants to the United States. 8

Many immigrants must rely not only on their hard work, extensive resources, and perseverance to succeed in America, but also a great deal of luck, which plays a significant role in an immigrant’s future in America. Prospective immigrants under the employment-based H-1B visa program know of this formidable game of luck all too well. Eligible H-1B applicants are the poster children for the American Dream. These applicants have the high-skill, intellect, and opportunity needed to succeed, yet most of them lack the one necessary element to work in the United States: luck. Subjecting these applicants to a feeble metric like luck fails to honor the fundamental ideal that “America is, and has always been, both a nation of immigrants and a nation of laws.” 9

The H-1B visa program (“H-1B”) allows foreign workers to live and reside in the United States through employer sponsorship. 10 On its face, it is a valuable program that stimulates America’s economy and protects U.S. workers. It seems to provide foreign workers with a clear and fair process to apply. However, the problem is that the H-1B requires highly specialized and qualified applicants and employers, yet it subjects these applicants and employers to an arbitrarily random selection process that makes no election for the applicants’ skills or qualifications. 11 Relying on the lottery to ultimately decide which foreign workers are admitted to work in the United States does not protect America or foreign workers because it subjects them both to a completely randomized process. Ultimately, the H-1B is stuck in a dichotomy that does not benefit U.S. businesses or foreign workers and fails at “the goal of building a fair, humane, and well-functioning

7 Although there are other employment based visas, the H-1B visa program is largely the only available option for skilled workers attempting to work in the United States because other employment-based visas are only available for smaller categories of people like professional athletes or temporary workers.
8 See id.
immigration system.” Accordingly, the H-1B needs to evolve with the changing needs of the nation and the fundamental notions of fairness for immigrants and U.S. businesses.

Although the entirety of the U.S. immigration system poses its own unique challenges, that discussion is outside the scope of this Note. This article will focus solely on the H-1B visa program and its random selection process. More specifically, this Note will discuss the reasons why the H-1B lottery is counterintuitive to the purpose of the visa program and why that disadvantages the foreign worker, employer applicants, and the American workforce. Before detailing the proposed solution, this Note first explains the background of the H-1B. This discussion includes the problems with the H category visa pre-1990 and the creation of the H-1B category visa in the 1990 Immigration and Nationality Act. In addition, Part I of this Note explains the H-1B visa’s petition and selection process.

Part II explains the original congressional intent of the H-1B. This part explains the dueling concerns of the visa to protect U.S. workers and gain a competitive advantage in the global economy. Part III addresses complaints with the H-1B, including the yearly cap set by Congress. Although the H-1B cap is controversial, this Note does not intend to partake in an argument about the merits of the cap. Instead, addressing the complaints with the H-1B cap helps to understand the flaws of the program and how some of those flaws may be fixed with solutions outside raising the yearly cap. Part III also discusses how large companies’ abuse of the H-1B leads to offshoring, which results in fewer opportunities for U.S. workers and prospective H-1B workers.

Part IV illustrates the numerous benefits of the H-1B, including innovation benefits that result in a more competitive U.S. market and economic benefits that create stronger U.S. companies and workers. Part V lays out merit-based immigration proposals and uses the proposed 2017 RAISE Act to show why the H-1B is an ideal candidate for a merit-based selection system. Part V also discusses the recent changes to the H-1B selection process and compares it to foreign merit-based points immigration systems. Part V concludes with addressing possible solutions to addressing merit and merging the H-1B’s employer driven model with a merit-based selection process.

---

Finally, this Note will end by addressing the benefits of merit-based selection systems, including why a new selection system will bolster the original intent of the H-1B, how the H-1B’s structure makes it an ideal candidate for a merit-based selection system, and potential problems with merit. Ultimately, this Note concludes with a discussion explaining how the H-1B addresses potential merit-based system problems.

I. OVERVIEW OF THE H-1B VISA PROGRAM

Over the past four years, hundreds of thousands of highly skilled and educated applicants registered for H-1B visas in each year. Among those applicants, only 65,000 are issued a visa, with an additional 20,000 issued specifically to qualified applicants with a master’s degree. This particular type of visa is intended for “foreign workers in specialty occupations, often referred to as 'high skilled’ occupations.” H-1B applications must show that the applicants possess highly specialized knowledge and at least a bachelor’s degree in the particular specialty. Primarily, these highly skilled applicants work in the science, technology, engineering, and math (“STEM”) fields. The Immigration and Nationality Act of 1990 (“INA”) revised the H-1B to narrow the scope of qualified applicants to only those employed in specialty occupations.

A. H Category Visa Before 1990

Prior to the 1990 INA, the H class visa consisted of three different categories, one of which was the H1 category. In the 1950’s, the H1 category had few labor protections for foreign workers and no cap on the number of visas allowed. The H1 category required applicants “to be of distinguished merit and ability” and to perform services “of an exceptional nature requiring such merit and ability.” Originally, the H1 visa also

14 Mauhan M. Zonoozy, America’s Stutter Towards H-1B Immigration Reform in America, 26 GEO. IMMIGR. L.J. 655, 655 (2012).
15 Id.
17 See Zonoozy, supra note 14, at 655.
20 B. Lindsay Lowell, Temporary Workers and Evolution of the Specialty H-1B Visa, 23 DEF. ALIEN 33, 36 (2000).
21 Jeronimides, supra note 19, at 369.
required applicants to establish an intent to return to their home country.22 By 1970, employers were allowed to hire for permanent positions.23 After the introduction of the H-1B visa in 1990, employees no longer had to maintain an intent to return home, so they could stay permanently.24 Congress sought to prevent abuse of the H1 Visa, while also protecting foreign workers with safeguards for their wages and working conditions.25 Additionally, the H1 category’s requirement that applicants be of distinguished merit and ability created an implied limit on the number of beneficiaries awarded the visa.26 Thus, the requirements prior to 1990 highlight the foundation of the H1 category visas: distinguished skill, qualifications, and ability.27

Debates around enacting the 1990 INA included a concern that the H1 program was no longer serving its original purpose.28 These debates included a growing issue that the merit-based requirements of the H1 program were becoming diluted as applicants no longer needed even a college degree for visa eligibility.29 As a result, many interest groups called for a change to rectify the dilution issue and clarify what distinguished merit actually meant.30 During the proposal of the 1990 INA, safeguarding the merit-based requirements of the H1 visa was crucial to honor Congress’s initial intent of the program and prevent a decrease in the number of jobs available to workers from the United States.31 Overall, “Congress’s intent in approaching the 1990 Act reforms consisted of a balancing of dual concerns: establishing a more efficient immigration system which is responsive to labor needs, while simultaneously according greater protection to both domestic and alien workers.”32

B. Inception of the H-1B Category

In the 1990 INA, the H1 program visas were split into two categories: H-1A and H-1B.33 As a result, those petitioning as entertainers, athletes, and other “prominent” people were placed into a new category.34 The new H-1B category limited applicants

22 Lowell, supra note 20.
23 Id.
24 Id.
25 Jeronimides, supra note 19, at 370.
26 See id.
27 See id.
28 See id. at 371.
29 See id.
30 See id.
31 See id. at 373–74.
32 Id. at 374.
33 See id. at 374 n.48.
34 See Paparelli & Patel, supra note 18, at 996.
to those in “specialty occupations,” which are occupations that demand “highly specialized knowledge and a bachelor’s degree or its equivalent.”35 Prior to the enactment of the 1990 INA, a focus intensified to create a distinction between applicants who truly have unique and specialized skills and applicants who are mere laborers.36 Although the INA changed the requirement from distinguished merit to specialty occupations, the logistic requirements remained virtually the same.37 This Act required H-1B petitioners to show their employment fit into the specialty occupation definition and established the controversial yearly 65,000 cap on the visa, not including the 20,000 allotted solely for master’s degree applicants.38

These new categorization standards and procedures aimed to tackle the complaints of the formerly known H category visas as well as provide greater safeguards for American workers.39 Among some of these safeguards, the labor condition application required H-1B petitioner employers to submit to the government assurances that the employer will provide the employee with wages greater than or equal to prevailing wages to protect the foreign and U.S. workers.40 Additionally, the INA required the Secretary of Labor to establish a process to receive and investigate complaints if employers failed to meet the requirements.41 Once again, the intent behind these safeguards and penalties was to protect both American and foreign workers while simultaneously “[r]ecognizing the hardship these new provisions might impose on the entry of foreign workers, and recognizing the value that foreign professionals might bring to American businesses.”42 The INA attempted to honor both of these intents through the petition and selection process.

35 See id.; see also H-1B Visa for Specialty Occupations, DOD Cooperative Research and Development Project Workers, and Fashion Models (Temporary Workers), U.S. CITIZENSHIP & IMMIGR. SERVS. [hereinafter H-1B Visa for Specialty Occupations], http://www.uscis.gov/forms/explore-my-options/h-1b-visa-for-specialty-occupations-dod-cooperative-research-and-development-project-workers-and [http://perma.cc/Q9JT-U3UL] (last updated May 18, 2022) (defining specialty occupation as a job that requires a bachelor’s degree or higher, is so complex that only a person with at least a bachelor’s degree can perform it, the employer typically requires a degree for the position, or the duties are so specific and complex that they are typically associated with attaining a bachelor’s degree or higher).
37 See Paparelli & Patel, supra note 18, at 998.
38 See Jeronimides, supra note 19, at 375; Zonoozy, supra note 14, at 655.
39 See Jeronimides, supra note 19, at 375.
41 See Jeronimides, supra note 19, at 377.
42 See id. at 378.
C. H-1B Petition and Selection Process

Before filing a petition, an applicant must ensure that they meet the requirements for the H-1B visa. To qualify for the H-1B, an employee must: (1) have an employer-employee relationship with the petitioning U.S. employer; (2) work at a job that qualifies as a specialty occupation; (3) have a job in a specialty field related to their field of study; (4) be paid the higher of the actual or prevailing wage in their field; and (5) have their employer file a labor condition application with the U.S. Department of Labor (DOL).  

If an applicant is eligible for the H-1B, their employer must follow a specific process before registering the employee for an H-1B petition because the H-1B is employer-driven. First, the employer must complete a labor condition application with the DOL to certify, among other things, that the H-1B worker will receive wages not lower than the prevailing or actual wages paid to similarly employed U.S. workers. 

Next, the employer must file a Form I-129 with the United States Citizenship and Immigration Services (“USCIS”) to verify certain information about the job, including employer and employee information. Employers bear the majority of the costs for an employee-applicant; employers pay the registration fee, filing fee, processing fee, and all other required costs and fees.

Once this petition is approved, the H-1B application is submitted and the USCIS determines if the applicant qualifies. When more petitions are received than the allotted 85,000 visas available for both master’s degree applicants and bachelor’s degree applicants, the USCIS uses a random selection process to fulfill the available visas. Before the 2020 amendment to the order of selection in the lottery, the USCIS randomly selected the 65,000 applicants followed by a random selection of the 20,000 master’s degree applicants. The two pools of applicants were kept separate. Now, the USCIS uses the lottery system to fulfill the 20,000 visas allotted for applicants with master’s degrees first. Then, the remaining unchosen master’s degree applicants and the entire

---

43 See H-1B Visa for Specialty Occupations, supra note 35.
49 Id.
How The H-1B Visa Lottery Prevents the American Dream

bachelor’s degree applicants are put into one pool and randomly selected to fulfill the remaining 65,000 open slots.50 Once an applicant is selected through the lottery, they are issued the visa for three years, with an option to extend for an additional three years.51

The quota for the H-1B visas has been met each year since the cap was set.52 Since 2014, the limit was reached within a week of the filing period.53 The USCIS is also required to issue the visas in the order in which the petitions are filed.54 With a new electronic registration process, the USCIS will open registration for fourteen days.55 At the close of the fourteen-day period, all properly submitted electronic registrations will be subject to the selection lottery.56 Under this process, all petitions filed within the period are considered filed at the same time “to ensure the fair and orderly administration of the numerical allocations.”57 However, the lottery process “brings anxiety to hundreds of thousands of employees and employers across the country, and the immigration attorneys and paralegals who assist them.”58 In part, this anxiety stems from the fact that specific rules and procedures for administering the H-1B visas can change each year and applicants have no way to increase their chances of selection.

In 2020, the USCIS announced a new rule to amend the H-1B registration process as well as change the order of petitions selected, resulting in a likely increase in the amount of master’s degree or higher applicants chosen to fill the 65,000 slots.59 This new rule switched the order in which petitioners were selected by requiring the USCIS to fulfill the 20,000 master’s degree or higher H-1B slots before fulfilling the 65,000 slots.60 This way, all of the unselected

50 See id.
51 See Karen Jensen, Barriers to H-1B Visa Sponsorship in the IT Consulting Industry: The Economic Incentive to Alter H-1B Policy, 35 FORDHAM INT’L L.J. 1027, 1033 (2012).
52 Dalecki, supra note 48, at 147.
53 Id. at 148.
54 See 8 U.S.C. § 1184(g)(3).
56 See id.
master’s degree applicants will fall into the pool of bachelor’s degree applicants, thus increasing a master’s degree applicant’s odds of being selected. The Department of Homeland Security (“DHS”) declared that this reversal in the selection process was consistent with the congressional intent behind the H-1B program to award H-1B visas to applicants who are the highest skilled.61 This is an attempt to lessen the randomized nature of the selection process; however, the lottery remains a gamble on percentages for employers and employees alike. The 2020 rule change merely makes the percentages smaller for bachelor’s degree applicants and higher for master’s degree applicants, but each must subject their fate to a random selection process.62

In 2022, the USCIS received over 308,000 registrations for the 65,000 available H-1B visas and the extra 20,000 master’s degree visas combined.63 In 2021, applicants had a thirty-two percent chance of being selected that season, without including the 6,800 petitions automatically set aside for workers from Chile and Singapore, which made the average applicant’s chances even lower.64 The exact percentages will change each year depending on the number of applicants and the qualifications of each, but overall, the chances for each applicant are quite low.65

II. INTENT OF THE H-1B PROGRAM

At the heart of the 1990 INA are multiple goals: facilitate the entry of immigrants to the United States to meet labor demands and protect opportunities of U.S. workers.66 This balancing act presents difficulties considering “immigrants created the nation we live in today.”67 By establishing the 65,000 cap on the H-1B visas, Congress intended to protect the competitive interests of the U.S. economy by encouraging market adjustments like better wages, more training, higher quality working conditions, and innovation.68 Before 1990, Democratic Congressman Bruce Morrison expressed concerns that the government gave out H-1

---

63 Id.
64 See id.
65 See id.
66 Jeronimides, supra note 19, at 380.
68 Lowell, supra note 20, at 36.
visas to foreign workers that were not highly skilled or necessary to the U.S. economy. Because of this, the 1990 INA attempted to mitigate the abuse of the system and curtail the visa program to reflect the system’s intent. Additionally, the visa cap was Congress’s answer to a growing fear that employers would exploit the weaker bargaining power of temporary foreign workers, thus displacing U.S. workers.

According to the DOL, the purpose of the H-1B is to “help employers who cannot otherwise obtain needed business skills and abilities from the U.S. workforce by authorizing the temporary employment of qualified individuals.” Furthermore, Congress’s intent for the H-1B was to aid American businesses by making them more competitive in the global economy while enhancing opportunities and wages of U.S. workers. When the 1990 INA took effect, it heightened the burdens on employers applying for H-1Bs on behalf of employees. In part, Congress enacted these burdens to address concerns that H-1B applicability and authorization would be overly generous, thus diluting the prestige and skill of the program. Members of the U.S. Senate expressed an immediate need for high skilled workers coupled with a need for greater innovation in STEM during a hearing before the Committee on the Judiciary concerning comprehensive immigration reform legislation. Additionally, one senator emphasized that STEM provisions need to be more responsive to the economic conditions of the country. That same senator raised concerns over the lack of visas granted for economic reasons.

---

69 How H-1B Visas Have Been Abused Since the Beginning, supra note 36.
70 See id.
73 See Jensen, supra note 51, at 1033 (arguing that Congress sought to balance the needs of U.S. workers with the need for high-skilled foreign labor while creating the H-1B program).
74 Paparelli & Patel, supra note 18, at 1001 (noting that the new law subjected employers to an elaborate enforcement mechanism and required petitioners to confirm that the employer would pay prevailing wages and provide prevailing working conditions).
75 See id.
76 See Comprehensive Immigration Reform Legislation: Hearing Before the Comm. on the Judiciary, 113th Cong. 13 (2013) (statement of Sen. Orrin Hatch) (explaining that an act that addresses the immediate short-term need for greater access to high-skilled workers as well as the long-term need to invest in America’s STEM education will provide the United States with the tools to compete in the global economy).
77 See id. at 22 (statement of Sen. Amy Klobuchar) (pointing to studies showing that the immigration policies that make it easier for professionals with special skills to work in the country result in more jobs for American workers).
78 See id. at 23 (statement of Sen. Amy Klobuchar) (“Under 10 percent of the core visas granted are for economic reasons . . . [and given the paramount need for economic growth that cuts across our ability to deal with all of our policy challenges . . . ].")
III. ISSUES WITH THE H-1B PROGRAM

Misuse of the H-1B program is not rare, especially in the IT industry. Exploitation of foreign workers obtaining H-1Bs poses significant problems, including the risk that employers will use the visa program for unskilled laborers, thus preventing the entry of numerous skillful applicants who might benefit the U.S. economy.79 However, this also exposes a significant issue for the H-1B program itself: it makes the extensive qualification, petition, and application process seemingly useless by allowing misuse that severely counteracts the very purpose of the program itself.

A. Issues with the H-1B Yearly Cap

For over twenty years, there has been a continuous, widespread debate over increasing the H-1B cap.80 The 1990 INA arbitrarily set the cap on H-1B visas as a precautionary means to reassure skeptics that the government would not dole out an unlimited amount of visas.81 Critics of the cap argue that the IT industry alone begs the need for more visas as it maintains such a large portion of STEM careers.82 In 1998, Congress enacted the American Competitiveness and Workforce Improvement Act (“ACWIA”), which raised the H-1B cap for five years to compete in the global market and prevent a labor shortage.83 After restoring the 65,000 cap in 2004, Congress added additional protections to the program, most notably a requirement that H-1B salaries and benefits were on par with U.S. workers in similar careers.84 Opponents of the cap also argue that the cap deters work from the U.S. market by compelling employers to offshore work when they cannot fulfill their needs domestically with either skilled foreign or U.S. employees.85

Additionally, these numbers may be severely disheartening to foreign students looking to study in the United States. With such meager chances of success, even after securing sought-after employment opportunities, foreign students have little incentive to waste their resources on American universities. Since

---

79 See Jensen, supra note 51, at 1040 (arguing that using the H-1B for unskilled laborers is especially concerning in light of the limited number of H-1B visas available each year).
80 See Lowell, supra note 20, at 38 (“Several bills have been considered that variously increase the numbers by at least 40 to 75 percent[] [and] introduce an unlimited H-1 for foreign graduates of U.S. schools] . . . .”).
81 Halliday, supra note 67, at 422.
82 Lowell, supra note 20, at 38–39 (noting that the IT sector accounts for approximately one-third of all STEM jobs).
83 Yaskey, supra note 71, at 900 (explaining that at this time, it was the first time the H-1B cap had been reached).
84 See id. at 900.
85 See id. at 901 (arguing that the cap “compels U.S. employers to move jobs outside the U.S. to new operations abroad.”).
“international students are becoming increasingly important to keep our classes full, our tuition revenue up, and our institutions thriving,” the ability, or lack thereof, of these students to stay in the United States is concerning for the American economy.86 These considerations become increasingly alarming in light of the fact that international students contribute around $39 billion to the national economy and support an estimated 400,000 jobs.87 However, it is becoming more difficult for these students to come to and remain in the United States.88

Furthermore, the H-1B cap is criticized for its negative impact on the U.S. economy and, specifically, American workers.89 Studies show that when companies are restricted in their ability to hire foreign workers due to the fact that foreign workers are unable to attain a visa through the H-1B lottery, American jobs are lost.90 For example, in areas that are not lucky enough to house the winners of the random H-1B lottery, careers in the computer industry decrease following the denials of H-1B visas.91 Whereas cities that are “lucky” enough to house more foreign workers experience an increase in careers for American workers in the computer industry.92

This research aligns with Microsoft’s experience. Microsoft previously vouched for these research findings, estimating that it adds around four employees to support every one H-1B hire.93 This demonstrates a positive correlation between H-1B approvals and American job growth. This also demonstrates the harsh truth that “employers of computer-related workers [do] not hire more natives when the foreign workers they intended to hire [are] denied H-1B visas.”94 The 65,000 H-1B cap has sustained challenges and the test of time somewhat due to the fear that foreign workers will displace American workers. However, the evidence overwhelmingly indicates that if substitution of an American worker for a foreign

---

87 Id. (noting the decline in international student enrollment in U.S. universities).
88 Id.
90 See id.
91 See id.
92 See id. (“The researchers find that foreign STEM workers stimulate new wage growth for both college-educated and non-college-educated Americans.”).
93 Id.
94 Id. (“More American jobs are lost when companies’ ability to hire foreign workers is restricted.”).
worker does occur, then it occurs among low-skilled laborers who, in theory, are not a part of the H-1B program.95

Increasing the H-1B cap is a meritorious argument and one that deserves strong consideration. But if Congress and the USCIS are unwilling to make this change, they should consider an alternative solution that gives applicants some agency over their outcomes rather than the current, almost hopeless, random thirty percent chance of success. Although this Note does not aim to explore arguments on raising the H-1B cap, proponents of raising the cap highlight that the problems with H-1B selection are deeper than a mere lack of availability. Overall, these problems indicate that restrictions on U.S. companies’ abilities to hire the right type of foreign workers eliminates jobs for the U.S. economy, Americans, and foreign workers petitioning to work in America.96 Arbitrary restrictions on the H-1B do not assist any of the goals of the visa program.

B. Misuse of the H-1B Program

Misuse of the H-1B program undermines the purpose of the visa and unmasks a critical need to alter the selection process. Large companies such as Disney, Southern California Edison, and Pfizer abuse the H-1B program by displacing American workers, hiring less experienced and lower-skilled foreign workers, paying the H-1B workers less, and outsourcing American work.97 Certain outsourcing firms account for nearly one-third of the available H-1B visas, yet these firms do not “use the H-1B visa as a way to alleviate a shortage of STEM-educated U.S. workers; they use it primarily to cut labor costs.”98 For example, foreign outsourcing firms utilizing the H-1B aim to maximize wage savings while outsourcing entry to mid-level tech jobs.99 Moreover, employers often take advantage of loopholes in wage requirements by hiring the foreign worker through a third-party service.100 In 2014, one study estimated that one-third of H-1B visas went to third-party outsourcing firms.101

---

95 See id. at 3.
96 See id. at 8.
98 See id. (exposing Southern California Edison’s tactics of replacing its original employees that made around $110,000 per year with overseas workers who were paid around $65,000 per year); see also Torres, supra note 44.
100 See Torres, supra note 44.
101 See id.
In addition, senators across party lines have worked to reveal the harm that these outsourcing firms bring to the H-1B, considering the fact that outsourcing firms are one of the biggest users of the visa program. In 2015, the Senate Judiciary Committee held a hearing to address the H-1B abuse, highlighting that “[a]lthough a small number of workers and students are brought in as the ‘best and brightest,’ most high-skill guest workers are here to fill ordinary tech jobs at lower wages.” Additionally, economists have reported significant research finding that these specific H-1B workers restrain wage growth, resulting in average employee earnings that are much less than they would be without the H-1B. This makes sense when a substantial number of the limited amount of visas are given to low skilled workers, who never actually fit the intended requirements for the visa. Complaints about H-1B abuse specifically target the IT industry. IT is the leading occupation field for the H-1B, yet many of these H-1B employees are low-skilled IT workers, making it difficult to believe that an actual shortage of these workers exists in America. As Nicole Torres explained:

Here’s a simplified way to explain how this plays out: Say you’re a big company with your own IT department. To reduce overhead, or to cut costs, or to increase efficiency, you decide to contract out (outsource) some or all of your IT work. So you hire an IT services firm to do that work on a temporary, as-needed basis. That firm sends workers, many of whom are on H-1B visas, to do those tasks. Sometimes, these contract workers supplement your IT staff; other times, you lay off your IT staff and the contractors effectively replace them.

The IT industry’s abuse of the program dilutes the talent, skill, and innovation the H-1B intended to bring to America. Because these IT firms successfully receive so many H-1Bs for unqualified workers, there are fewer visas available for in-demand, highly skilled applicants. Since the H-1B program is not working the way it was intended, the U.S. economy is...
negatively impacted in the global market. Structurally, the H-1B created a “one-size-fits-all” system, but this lottery system does not combat the companies’ misuse. Instead, the lottery switches the focus from an applicant’s skill and the company’s need for the foreign worker to reward companies that have learned how to game the system, benefiting those that have “learned how to take advantage of loopholes.” Thus, an arbitrary lottery is not responsive to the misuse of the program, nor is it responsive to sorting through the enormous pools of applicants.

IV. BENEFITS OF THE H-1B PROGRAM

Despite complaints, the H-1B creates important benefits for the country. Even with the abuse of the H-1B, the numbers do not lie. For example, the Department of Labor received 269 complaints on abuse of the H-1B in 2001, yet only fifty-four cases were found in violation. Furthermore, “[c]onsidering the number of H-1B workers employed in the United States, the number of complaints is still relatively small.” Moreover, the H-1B has made the United States “the beneficiary” of the most talented and acclaimed scientists, engineers, economists, and other professionals in the world.

A. Innovation Benefits

The H-1B provides foreign students studying in the United States with one of the only avenues to gain employment and remain in the United States. Without the H-1B, se talented students will have to take their talent and potential innovation elsewhere, which means U.S. universities’ resources would be wasted as well. Moreover, scientists and engineers make up around one-fourth of total productivity growth in the United States, and the majority of H-1B applicants work in STEM fields. Also, the “one thing that has helped maintain [the United States’s] technological leadership is innovation and technical research, and immigration has helped [the United States] do

---

109 See id.
110 Daniel Aobdia, Companies Want to Hire the Best Employees. Can Changes to the H-1B Visa Program Help?, KELLOGG INSIGHT (Feb. 6, 2017), http://insight.kellogg.northwestern.edu/article/how-to-revamp-the-visa-program-for-highly-skilled-workers [http://perma.cc/7KME-42T7].
111 Id.
112 Halliday, supra note 67, at 430.
113 Id.
114 Id. at 426.
116 See Thompson, supra note 104.
117 Gill, supra note 115, at 244.
How The H-1B Visa Lottery Prevents the American Dream

The H-1B aids American employers because it allows them to fill in the gaps in the domestic workforce. In turn, this enables U.S. employers to compete in the global economy.

Without the H-1B, the United States cannot meet its demand for workers, and companies will have to ship their operations offshore. Furthermore, Congress arguably implemented proper safeguards to reduce H-1B abuse and to protect U.S. workers. This extensive application process makes hiring U.S. employees easier and more accessible, increasing the likelihood that qualified U.S. workers will fill those positions first.

B. Economic Benefits

The H-1B provides numerous economic benefits. First, the H-1B stimulates the economy by increasing opportunities within the United States. Large tech companies have emphasized that a lack of H-1B availability causes them to shift work overseas, or even as close as Canada. Offshoring harms the U.S. economy and U.S. workers alike because it takes the opportunities, profits, and innovation out of the country. Thus, the availability of H-1B workers "creates a positive externality for the companies that hire them and the economy because the workers create new products, and in some instances, new sectors of an industry, which creates opportunities for other workers." While some opponents of the H-1B argue that it leads to offshoring, the result is the opposite; the program may prevent high-skilled workers from taking their skills elsewhere and instead allows the U.S. job market to grow for foreign and domestic workers. The problem arises when corporations cannot find either qualified American or qualified foreign workers to fill needed positions because they are forced to offshore positions elsewhere or when employers are not allowed to hire qualified foreign workers.

Additionally, H-1B workers contribute to the American economy through taxes. On average, H-1B workers earn

---

118 Torres, supra note 106.
119 Halliday, supra note 67, at 422.
120 Yaskey, supra note 71, at 899.
121 See Halliday, supra note 67, at 426.
122 See id. at 428.
123 See id. at 428–29.
124 See Yaskey, supra note 71, at 901
125 Id.; see also Halliday, supra note 67, at 432–33.
126 Yaskey, supra note 71, at 901.
127 See Halliday, supra note 67, at 432.
128 See id. at 432–33.
between $100,000 and $120,000 and pay taxes between twenty
and forty percent.130 It is estimated that the current three million
H-1B visa holders contribute, on average, $27.1 billion to
American programs that they cannot benefit from, while also
contributing around $76.7 billion annually to the U.S. economy.131

The benefits of the H-1B have the potential to greatly outweigh
its criticisms. Nevertheless, the random selection process does not
reinforce any of the benefits the program creates. Arguably, the
random selection does the opposite by thwarting the program’s
ideals, purposes, and benefits with a lottery selection. In addition,
the USCIS receives criticism for the random selection process and
the lack of information available on the process’s
implementation.132 Critics brought suit in 2016 against the USCIS
intending “to pry open that box and let the American public and
those most directly affected see how the lottery system works from
start to finish, and to learn whether the system is operating fairly
and all the numbers are being used as the law provides.”133
Moreover, petitioners of the H-1B sued the USCIS over the
“arbitrary and capricious” randomized lottery selection system.134
Employers argued against the selection process, “as it results in a
potentially never ending game of chance for petitions filed . . . with
some unlucky individuals trying and failing each year to obtain a
quota number, while some lucky lottery winners obtain a visa
number in the very first year a petition is filed on their behalf.”135

Clearly, employer sponsors and foreign workers are fed up
with remaining in the dark over the selection process that
determines their futures. The focus of these lawsuits is to find a
solution that gives applicant employees and employers a fair and
orderly system.136 Many employers and foreign workers know that
raising the H-1B cap is likely a losing battle; all they are calling
for is a selection process that is transparent, logical, and beneficial
to everyone involved in the program. The government went
through great pains and deliberations to create a sophisticated

---

130 Id.
131 Frank Gogol, The True Economic Impact of H1B Visa Holders, STILT (Apr. 28,
[http://perma.cc/KCM2-4TJH].
132 See Natalie Olivo, FOIA Suit Seeks to Peek into ‘Black Box’ of H-1B Lottery, LAW360
(May 23, 2016), http://www.law360.com/articles/799213/foia-suit-seeks-to-peek-into-black-
133 Id.
134 Suevon Lee, Rejected H-1B Visa Applicants Slam ‘Arbitrary’ Lottery, LAW360 (June
2, 2016), http://www.law360.com/articles/802633/rejected-h-1b-visa-applicants-slam-
135 Id. (quoting Complaint, Tenrec, Inc. et al. v. U.S. Citizenship & Immigration Servs.
et al., No. 3:16-cv-00995 (D. Or. 2016)).
136 See id.
V. Merit-Based Solution

In order to accurately honor the intent of the H-1B and ensure that the annual 85,000 visas are granted to qualified applicants, the random selection process needs to recognize various levels of merit between applicants. Merit-based means qualifying applicants based on various levels of skill, experience, and capability. Numerous other countries have implemented merit-based immigration systems, and many American politicians have proposed the same. Merit-based solutions include wage-based selection, points-based selection, and employer-driven selections models.

A. Wage-Based Selection

Rather than using a random selection process, a wage-based selection system prioritizes the highest salary levels. In January 2021, the USCIS announced a final rule that would change the H-1B lottery selection process. The aim was to shift to a wage-based selection system. This new wage-based selection system is set to take effect in the 2023 selection process. The USCIS modified the selection process in hopes that it would “incentivize employers to offer higher salaries, and/or petition for higher-skilled positions, and establish a more certain path for businesses to achieve personnel needs and remain globally competitive.” However, this modification of the selection system is just a start, as wage-levels do not address all the problems nor enhance the benefits of the H-1B. Moreover, prioritizing wage fails to acknowledge numerous other important factors that play a role in the success of foreign workers and the competitiveness of the

---

138 See id.
140 See What H-1B Lottery Results Says About Odds?, supra note 62.
142 Id.
United States in the global economy. Focusing on wage alone negatively impacts foreign talent coming straight out of college.

Wage prioritization fails to appreciate the skill and innovation entry-level workers may bring to the American workforce. Entry-level workers may still be viable H-1B applicants because entry-level does not correlate to low-skill, and these applicants may still possess the required education in a specialty occupation. The United States needs foreign students because of their positive financial impact on American universities, as “international students are a vital source of revenue.” Yet, these students will likely be left with few options to remain in the United States past their education if the H-1B focuses solely on wage levels. Many international students looking to study in the United States will now likely look to the U.S. neighbor, Canada. Employment laws in Canada aid international students, while international students in the United States must hope their luck runs strong in the H-1B lottery. Furthermore, prioritizing H-1Bs based on wage could “harm start-ups, small businesses, nonprofits, and rural businesses that do not have the resources to offer the highest wages.” These businesses have the need for H-1B workers, but prioritizing based on wage will likely wipe out an entire market that exists for skilled workers, replacing them with large tech companies that can afford to keep up with this new rule.

A mere change of the order in which petitions are “randomly” selected demonstrates a much larger need: merit-based decision making for a distinguished visa. Many have argued, lobbied, and even begged for a change within the H-1B; some have called for an increase in the annual visas available, while others have criticized the USCIS’s adjudicatory actions or lack thereof. Replacing the now semi-random selection process with a complete merit-based selection process would meet the underlying concerns coming from both sides of H-1B debates. Proposals for a merit-based immigration system are not new; however, this proposal focuses

---

145 Id.
147 Id.
148 See Jeronimides, supra note 19 at 385–86.
specifically on a merit-based selection system in replacement of the H-1B random selection process.

B. Points-Based Selection System

Many countries have successfully implemented merit-based selection processes through a points system. Points-based selection systems reflect merit because they allocate points based on valued skills, education, and experience. Prior proposals for merit-based immigration usually mimic the merit-based approaches used in Canada, Australia, and New Zealand. Although Canada, Australia, and New Zealand are the most commonly cited examples of merit-based immigration systems, almost every EU country, as well as many others, use some form of a merit-based system. In 2017, the Reforming American Immigration for Strong Employment (“RAISE”) Act was introduced to the Senate as an amendment to the 1990 INA, advocating for a point-based immigration system. Under this system, an applicant would receive more points based on their level of degree. Currently, employment-based immigrants may enter the U.S. in the following five different ways:

1. “persons of extraordinary ability” who can document high-level accomplishments in their field, a tough standard for most workers to meet, particularly early in their careers;
2. professionals with advanced graduate degrees or exceptional ability;
3. professionals, skilled workers (those with at least a two-year college degree), and unskilled workers (for whom just 5,000 green cards are allocated yearly);
4. certain special immigrants who meet U.S. national interests; and
5. immigrant investors, who invest at least $500,000.

Under this system, applicants who qualify for the H-1B fit into at least categories (2) and (3), with the possibility of fitting into more. Under the RAISE Act, applicants with the following eligibility criteria are given higher priority (more points): 26-30

---

154 US Merit Based Immigration, supra note 151.
years old, holds a U.S. STEM degree, proficient in English, extraordinary achievements, job offers and salary, and the applicant’s likelihood to invest in the U.S. economy. The points-based selection proposed under the RAISE Act emphasizes many of the same skills and characteristics that U.S. employment-based visas already require.

The RAISE Act proposed this points-based system as a response to an influx of low-skilled labor, resulting in lower wages for foreign workers—as well as American workers. Ultimately, the RAISE Act failed, as it sought a complete overhaul of the American immigration system, including non-employment visas. Additionally, the RAISE Act proposed eliminating any existing employment categories, remodeling the entire system already set in place. Other drawbacks of the RAISE Act included the possibility that an applicant’s points could decrease if their spouse immigrated with them. However, the RAISE Act also proposed a recurring scheduled re-evaluation of the points system to ensure that the points awarded continued to meet the needs of the United States.

At these RAISE re-evaluation meetings, proposed changes in points would be based on concerns in four main areas: “(1) increasing the United States’ per capita growth in gross domestic product; (2) enhancing the chances of financial success for point-based immigrants; (3) cultivating the ‘fiscal health’ of the United States; and (4) maintaining or increasing wages for domestic workers.”

H-1B applicants would also fit into most of the categories proposed under a points-based system, as they also must have a job offering prevailing wages that requires at least a bachelor’s degree in a specialized field (usually STEM). Furthermore, the majority of petitioners are between 25 and 34 years old. On its own, the H-1B is already somewhat of a merit-based program, yet it subjects its applicants to a meager chance of success within a completely

156 See US Merit Based Immigration, supra note 151.
158 See id. at 483.
159 See id. at 484.
160 See id.
161 See id. at 485.
162 Id.
164 Id.
How The H-1B Visa Lottery Prevents the American Dream

arbitrary selection process. Given that the number of H-1B applicants outnumber the H-1Bs granted by an extraordinary amount, it is unfair to assess a candidate who merely meets the criteria in the same manner as a candidate that contributes to U.S. industries and the American economy in a more meaningful way.

C. Foreign Points-Based Immigration Models

Some countries, such as Canada, use a points-based immigration system to favor factors including formal education and language proficiency. Alternatively, Australia’s system favors specific skills on an as-needed basis and contributions to public finances. Additionally, New Zealand uses a points-based program that grants skilled workers more points when they have already gained relevant work experience in the country. Moreover, these countries adjust the points system based on the market needs of that year. The success of these merit-based programs is reflected in the fact that talented international students are looking to study in other English-speaking countries like Australia and Canada. Aside from Australia, Canada, and New Zealand, a number of other countries employ a form of merit-based immigration—including Austria, Denmark, Germany, Japan, South Korea, and the United Kingdom.

Canada has become a leading country for international students, partially due to the fact that the Canadian visa process is simpler and shorter. In Canada, international students have a higher chance of obtaining permanent residency compared to the United States. Canada is not the only country gaining competitive status in the race for higher education international students; Australia has promoted its educational system and gained nearly $16 billion in income for that industry. In contrast to the RAISE Act proposals, Canada’s points-based

167 Id. at 4.
168 See id. at 5.
169 See Krislov, supra note 86.
170 See Parsons, supra note 157, at 485.
172 See Hegarty, supra note 144, at 229.
system applies only to its designated economic class of visas.\textsuperscript{175} Comparatively, Australia’s system is separated into two programs: a Migration Program and a Humanitarian Program.\textsuperscript{176} The Migration Program deals with skilled foreign workers, similar to the U.S. H-1B.\textsuperscript{177} Countries like Australia and Canada that have adopted points-based systems, were able to create immigration systems that balance a need for skills-based workers, family-based immigration, and low-skilled worker visas. These systems balance those interests by creating different selection processes based on the specific visa.\textsuperscript{178} This is specifically where the RAISE Act fell flat—proposing to eliminate different family immigration categories and replacing them completely with the skills-based points system.\textsuperscript{179}

Some of the foreign merit-based systems focus on employer demand and potential economic benefits of select visa holders.\textsuperscript{180} Specifically, Australia has tailored its program to prioritize employer-specific needs.\textsuperscript{181} Canada’s program uses an online service to assess an applicant’s initial eligibility, and once an applicant is deemed eligible, they are awarded points based on different factors of an adjustable scale.\textsuperscript{182} These factors include “human capital, the transferability of [an applicant’s] skills and their ability to integrate successfully.”\textsuperscript{183} In many of these skills-based points systems, foreign workers give general information regarding their qualifications, education, experience, and identity, while only the highest-ranked applicants are given an opportunity to apply.\textsuperscript{184} Canada’s system has received praise “both for its transparency, since it involves clear and objective selection criteria, and its flexibility, as policymakers can adjust the system’s criteria and distribution of points in response to changing estimates of labor demand.”\textsuperscript{185} Many countries using some form of a merit-based points system have to find a balance between economic qualifications and human capital.\textsuperscript{186} This balance of economic qualifications includes weighing factors like education, experience, and job offers against human capital

\textsuperscript{175} Parsons, supra note 157, at 488.
\textsuperscript{176} Id. at 490.
\textsuperscript{177} See id.
\textsuperscript{178} See id. at 491.
\textsuperscript{179} See id. at 483.
\textsuperscript{181} See id.
\textsuperscript{182} See id.
\textsuperscript{183} Id.
\textsuperscript{184} See Parsons, supra note 157, at 492.
\textsuperscript{185} See Maurer, supra note 180.
\textsuperscript{186} See Merit Based Immigration and How it’s Measured, supra note 152.
factors, such as needs of the industry. Without this balancing act, countries face the problem of highly qualified immigrants working in jobs that they are overqualified for. Thus, how a system defines merit is crucial for its success.

D. Employer-Driven Models

Other proposals to improve the H-1B include replacing the lottery with an auction, so supply and demand dictate the value of a worker. This would mean employers would bid on employees, and employers with the highest bids “win” the employee. However, an auction-like solution fails to appreciate the importance of foreign workers having agency over their employment. An auction would take a visa that is already employer-focused and completely eliminate the employee from the process. Furthermore, this selection process leaves foreign workers “at the mercy of th[e] company” that successfully bids on them. The employer-driven model of the H-1B benefits the United States because employers have better qualifications to choose workers best suited for the company and the country. Overall, the high-skilled H-1B workers tend to earn more, produce more patents, and outperform workers who enter on green cards. Therefore, any changes in the H-1B selection process should not diminish or significantly alter the benefits of this model. With an employer-driven model, employers guide the beginning of the H-1B process, as they determine what applicants get in the door initially; however, a successful model will honor the qualifications of the employer along with the agency of the immigrant worker.

E. Measuring Merit Depending on the Country’s Immigration System

Other countries’ merit-based immigration systems confirm that every program does not function the same, and countries need adaptable systems to fit their specific needs. For example, the United Kingdom’s system grants points for specific skills, professions, salaries, and qualifications; moreover, applicants

---

187 See id.
188 See id.
189 See id.
190 See id.
191 Id.
must have a job offer from an approved employer. Canada’s program awards points for English or French language skills, education, and relevant work experience, while an employment offer merely grants an applicant additional points. Australia’s points-based system prioritizes skilled employment experience and education by requiring that all applicants have their skills assessed. New Zealand’s system closely imitates Australia’s but places more focus on the market needs of the country by requiring that an applicant has skills the country currently needs. Each of these systems has its own benefits and deficiencies, but each system also shows a global development in the merit-based immigration arena. The H-1B’s history recognizes that the program “endeavored to facilitate the ability of American businesses to compete in the global economy.” While other countries are leveling up to create employment-based visas that entice skilled foreign workers with clear guidelines and formulas for success, America’s H-1B falls flat with its randomized lottery. Some of these countries enacted points-based systems as recently as 2020, a sign that America needs to follow the global trends to remain a top global competitor.

VI. BENEFITS OF MERIT-BASED PROGRAMS

A merit-based system enables a country to select immigrants more likely to earn higher amounts and contribute more to the government. Canada’s merit-based system shows that these high-skilled workers admitted under a points system have better qualifications, higher education, more robust employment rates, and are more likely to make higher contributions to the government. Similarly, the U.S. labor market benefits when workers perform jobs that meet their specializations. Much of the criticism surrounding the H-1B includes mistreatment of the requirement that workers coming through the program must be highly-skilled. A merit-based points selection process clarifies what “high-skilled” means and redefines the concept of merit.

---


194 Id.

195 Id.

196 See id.

197 Paparelli & Patel, supra note 18, at 997.


199 Hunt, supra note 192.

200 Id.

201 Id.

202 See Palagashvili & O’Connor, supra note 89, at 2.

203 See Aobdia, supra note 110.
This creates a more streamlined process because it refines the necessary requirements and raises the bar for selection. A higher bar for selection would allow the most high-skilled H-1B applicants to improve their chances of success from the meager thirty percent chance they currently face.\textsuperscript{204} The points system would allow applicants to assess where they fall on the spectrum and give them metrics to improve upon to set themselves apart from other applicants.

A. Merit-Based Selection Process Bolsters Intent of the H-1B Program

The requirements of the H-1B itself, as opposed to the requirements for the visa selection process, are already merit-based. The eligibility requirements in the proposed 2017 RAISE Act include much of what is already required for any applicant under the H-1B. Much like the proposal under the RAISE Act, the H-1B already prioritizes those with higher level degrees by running the lottery first on the 20,000 master's degree applicant pool before conducting the lottery for the remaining 65,000.\textsuperscript{205}

Currently, the H-1B is employer focused as it enables employers to seek out and apply for foreign workers.\textsuperscript{206} However, employers and applicants lose all their agency in the selection process for the H-1B.\textsuperscript{207} Certainly, employers and applicants cannot have complete control over the selection process because the government ultimately regulates entry for foreign workers. However, giving employers greater agency in the selection process reinforces the idea that employers are best suited to understand the market needs and pick employees who will fulfill those needs. Replacing the lottery with a merit-based selection process also incentivizes employers to find applicants who are best suited for the position and possess important skills and qualifications, instead of scouring the foreign worker market for cheaper labor.

Furthermore, prioritizing applicants with truly valuable skills disincentivizes abuse of the program.\textsuperscript{208} Employers who previously found loopholes to use the H-1B as a means to hire lower-skilled workers and pay them less will not be granted competitive visas because more desirable, higher-skilled applicants will have more points. Moreover, adding guidelines to the selection process “would reduce stress and uncertainty for the most desirable foreign

\begin{footnotes}
\footnotetext[204]{See What H-1B Lottery Results Says About Odds!, supra note 62.}
\footnotetext[205]{See Zonoozy, supra note 14, at 655.}
\footnotetext[206]{See Hunt, supra note 192.}
\footnotetext[207]{Nita Nicole Upadhye, H-1B Visa Lottery (How Does it Work?), NNU IMMIGRATION (Apr. 1, 2021), http://www.nnuimmigration.com/h1b-visa-lottery/ [http://perma.cc/5UM4-M7YM ].}
\footnotetext[208]{See id.}
\end{footnotes}
applicants, many of whom graduate from top-tier MBA, PhD, and engineering programs hoping to stay in the country, as well as for the companies that wish to invest in these employees.”209 If the intent of the program is to aid the competition of American businesses in the global economy and enhance the opportunities and wages of U.S. workers, then the selection process should reflect that intent. A random selection process certainly does not enhance opportunities on either side of the visa.

B. H-1B's Structure is an Ideal Candidate for Merit-Based Selection

Comparisons of other countries’ merit-based immigration systems highlight why the H-1B is an ideal candidate for a merit-based selection system. The United Kingdom, for example, requires that skilled foreign workers have employer sponsorship, whereas Australia admits applicants based on their points earned.210 Debates around points-based systems have heavily focused on the advantages and disadvantages of employer sponsorship. The majority of countries that use points-based systems differ in their requirements that an applicant is sponsored by an employer.211 Points-based systems often do not require job offers, so there is no guarantee that a “high points” foreign applicant is employable in the necessary market or that the host country needs that type of work.212 The H-1B circumvents this problem because the visa’s petition requirements are employer-driven and require that the foreign worker has employer sponsorship. Without this employer-driven system, points-based systems must rely on the government to decide what skills the country values.213 Employers are better suited than the government to assess the needs of the market and their companies and to evaluate what type of worker will best fit their needs.

Employers rely on this visa to employ foreign workers because it provides the most streamlined process for an applicant to become a permanent resident.214 This means employers utilize H-1Bs with long-term goals in mind, keeping high-skilled foreign workers in the

209 Id.
211 See id. at 7.
212 See id.
213 See id. at 5.
U.S. market. The government is not equipped to make these decisions for employers, and it removes decision-making skills from employers when they use the lottery to grant H-1Bs. The government does not currently incentivize employers to choose those who will best contribute to the market because the H-1B selection process does not differentiate between applicants. As of now, the H-1B selection process differentiates applicants on two measures: salary and degree level. Regarding degree level, applicants with a master's degree or higher have preferential selection because the lottery runs first for applicants in this pool. The applicants not chosen receive an additional opportunity in the lottery for applicants with at least a bachelor’s degree.

C. Measuring Merit

A merit-based solution for the H-1B selection process should not be a complete overhaul of the existing system. Currently, highly educated immigrants perform better in the United States than in Canada, and the H-1B is already tailored to these immigrants. Overall, a points system should decide on desirable characteristics, assign each characteristic a different point value, and remain flexible enough to consider market needs. A successful system would honor the value and needs of foreign workers, the American workforce, and U.S. employers. Thus, a complete restoration of the H-1B is unnecessary because parts of the system already work well. The following should be considered when crafting the ideal merit-based system:

The ideal merit-based system is not simply points-based . . . . “It is a dynamic, agile system that is actively operated by a hands-on government and carefully calibrates the profile of the future workforce. Many countries with points-based systems have been moving away from just selecting people on the basis of human capital attributes and giving more value to an employment offer.” . . . “Ultimately keeping employer sponsorship and/or heavily weighting a job offer under the point system will be extremely important, along with clearing the green card backlogs.”

The H-1B is structured to bring in high-skilled applicants with an education, related job offers, and valuable skills. This program already does what many merit-based immigration systems aim to do, so it makes sense to slightly alter the selection process to stay consistent with these goals.

215 See id.
216 See What H-1B Lottery Results Says About Odds?, supra note 62.
217 See Hunt, supra note 192.
218 Maurer, supra note 181.
D. Problems with Merit-Based Systems

Critics of merit-based immigration argue that points-based selection systems ignore the potential of an “immigrant superstar,” the “entrepreneur with a billion-dollar idea [who] may generate more value than thousands of dependable, hardworking, high-skilled workers.”\(^{219}\) Because merit-based points systems value “reliable” metrics like education, experience, and career opportunities, these systems risk undervaluing those who could theoretically generate more value from riskier endeavors like startups and entrepreneurships.\(^{220}\) In theory, a successful start-up could offer more employment opportunities for foreign and U.S. workers, generate more profit than a “traditional” H-1B worker, and offer more to the U.S. economy overall.\(^{221}\) However, these problems were addressed under the RAISE Act proposal—which proposed a system that accounted for the U.S.’ gross domestic product, the chances of financial success for the admitted immigrant, the overall fiscal health of the United States, and increasing wages of the whole workforce.\(^{222}\) Specifically, the system under the RAISE Act would award points for things like an investment in a business, an extraordinary achievement—and playing an active role in management of a new commercial enterprise.\(^{223}\) Additionally, the current H-1B random selection process does not offer any better of a chance for foreign-born startup founders, as they may be subject to the lottery year after year without success and with no chance of better odds the next year.

However, some current skills-based points systems used in Australia and Canada do not recognize the importance of employment opportunities, as they decrease the number of points offered to an applicant for a job offer.\(^{224}\) This creates a system that values skill, without considering opportunity—resulting in high-skilled, foreign workers competing for unskilled jobs, which in turn creates problems for the workers and the economy.\(^{225}\) Thus, systems that fail to recognize the significance of a current employment opportunity lead to a problem called “brain waste”—the problem of foreign-trained doctors driving taxis.”\(^{226}\)

\(^{220}\) See id.
\(^{221}\) See id.
\(^{222}\) Parsons, supra note 157, at 485.
\(^{224}\) Parsons, supra note 157, at 492.
\(^{225}\) Id. at 492–93.
\(^{226}\) Gelatt & Neufeld, supra note 219.
Therefore, a better merit-based system would understand and utilize the importance of employers choosing their desired workers.\textsuperscript{227} The RAISE Act also failed to accurately appreciate that an employer-driven immigration model allows employers to select the skilled workers they need.\textsuperscript{228} The RAISE Act proposed taking the power away from employers, who are best suited to understand the needs of the market, whereas the H-1B program already recognizes this importance.

The demise of the RAISE Act included concerns that the agricultural, hospitality, and manufacturing industries would face severe labor shortages because of the points-based selection system.\textsuperscript{229} Overall, the RAISE Act faced criticism, and ultimately failure, because its points system worked at the expense of lower skilled, laborer immigrants—cutting off family-based categories and significantly reducing the number of immigrants admitted to the United States.\textsuperscript{230} Nevertheless, the merit-based points system proposed by the RAISE Act was not the issue. The issue was that the visa systems the RAISE Act proposed changing did not need to address merit. However, the H-1B is already structured to address merit, making it an ideal candidate to enact a similar merit-based points system proposed under the RAISE Act. Unlike the RAISE Act’s proposal, this is a narrow solution to a very specific problem.

E. H-1B Program Answers the Problems of Merit in its Structure

Although the U.S. H-1B selection process does not currently use a merit-based points system, the program already avoids the issue of “brain waste” because the visa is employer-driven; a U.S. job offer is a prerequisite to apply. Furthermore, an applicant’s job offer must be one that qualifies as “highly skilled.”\textsuperscript{231} Therefore, implementing a merit-based system in lieu of the current lottery would not risk the issues of “brain waste” other merit-based systems face because the H-1B requirements have already accounted for this issue. The H-1B requirements already consider the fact that employers are in the best position to recognize what talent is best suited in their lines of work, so the selection process should mirror “[a] broader, more flexible merit-based immigration stream [that] would supply employers with needed workers, support U.S. workforce growth amid an aging population and

\begin{itemize}
  \item \textsuperscript{227} Id.
  \item \textsuperscript{228} Koons, \textit{supra} note 223, at 146–47.
  \item \textsuperscript{229} Id. at 149.
  \item \textsuperscript{231} See 8 U.S.C. § 1184(c)(6)(A)(ii).
\end{itemize}
bolster overall U.S. economic growth.”232 Currently, the random selection process fails to achieve either of those goals. At this time, employers and employee applicants are subjected to strenuous requirements to show merit, but then must leave their fate up to a random selection process that does not provide employee applicants with higher chances at staying in the United States.

Replacing the H-1B lottery with a merit-based points selection process will honor the original intent of the program, retain the employer-driven model of the visa, and give employers and applicants greater agency over the process—creating an overall more equitable and valuable system. Maintaining the employer-led and demand-driven structure of the H-1B is important to the success of the program because it facilitates economic growth and competitiveness.233 A points-based and skills-focused model, without components of a demand-driven model, runs the risk of high-skilled immigrants unable to find jobs that fit their prominent skill sets.234 Without focusing on demand, the value of a high-skilled immigrant working in the United States is not accurately reflected onto the American economy.235 The H-1B is demand-driven because the employers are in the driver’s seat; foreign workers cannot work in the United States under an H-1B without employer sponsorship.236 Data comparisons between the U.S. demand-driven model and Canada’s point-based model demonstrates that points-based systems are more effective in obtaining higher educated foreign workers.237 Coincidentally, this is precisely the intent of the H-1B. However, Canada’s system fails to honor these high-skilled workers’ skills because many of them end up working in jobs they are substantially overqualified for.238 Thus, a merit-based selection system creates a best-of-both-worlds situation because it maintains the employer-driven demand model for H-1B eligibility while implementing a “best and brightest” selection system.

Countries like Canada, Australia, and New Zealand that have adopted merit-based immigration systems prove that merit-based immigration is not a one-size-fits-all. For instance, Australia’s system is a hybrid system, combining a points-based system with

232 Gelatt & Neufeld, supra note 219.
234 See id.
235 See id.
236 See id.
237 See id. at 31–34.
238 See id. at 34.
How The H-1B Visa Lottery Prevents the American Dream

a demand-based employer model. Although these countries provide a framework for what a merit-based system could look like, most of these countries have amended their systems to combine the need for employer-driven demand and pure points. The U.S. H-1B program should follow suit.

CONCLUSION

The H-1B should not be another contributor to the failed American Dream. The American Dream fills skilled H-1B applicants with a false sense of hope that their hard work, perseverance, and dedication to honoring the system will end in success. For many H-1B applicants, around 70% per year, this is simply not true. H-1B applicants represent most of what the government says the United States wants and needs, yet the government ties the hands of all prospective applicants and expects them to accept the random lottery as the holder of their fate. As Oscar Handlin said, “the immigrants were American history,” but they are also the future. America needs to honor this with a H-1B selection process that reflects the merit, influence, and importance of the applicants.


240 See Hunt, supra note 192.

241 Vitello, supra note 1.
Challenging Subpar Servitudes

Samantha Kuo*

INTRODUCTION ..............................................................................................................298
I. HOUSING IN CALIFORNIA: PAST AND PRESENT ........................................301
II. GOLF COURSES CONSERVATION EASEMENTS: AN UNPRODUCTIVE USE OF LAND .................................................................304
   A. The Declining Popularity of Golf .................................................................304
   B. Golf Courses Do Not Serve a Legitimate Conservation Easement Purpose ................................................................................................................306
      1. Golf Courses Do Not Serve the Statutory Intent of Conservation Easements .........................................................................................306
      2. Golf Courses Do Not Serve the Environmental Objective of Conservation Easements .................................................................308
   C. Replacing Golf Courses with Housing Is a Productive Use of Land ...........................................................................................................310
III. THE PERPETUITY FEATURE OF CONSERVATION EASEMENTS: THE CONS OUTWEIGH THE PROS .................................................................311
   A. All Conservation Easements Are Actually or Constructively Perpetual ........................................................................................................312
   B. The Perpetuity Feature of Conservation Easements Serves Only a Few Purposes .................................................................314
   C. Perpetual Conservation Easements are Harmful Because They are Ineffective and Bind Future Generations to Outdated Science and Cultural Values ...........................................................................................................315
IV. PROPOSAL ..............................................................................................................320

* J.D. Candidate 2023, Chapman University Dale E. Fowler School of Law. Thank you to Professor Kenneth Stahl for sharing his expertise in land use and for guiding me through my first major research paper. Thank you to Professor Sherry Leysen for assisting in my research by contributing her priceless research knowledge and her optimism. Thank you to the staff of the Chapman Law Review for the time and energy they spent improving this Note. Finally, thank you to my family, friends, and partner for making law school possible and for making life enjoyable.
A. The California Legislature Should Amend the 1979, 1974, and 1969 Acts ................................................ 320
B. The Courts Should Apply a Balancing Test to Terminate Conservation Easements ...................... 321
   1. Weighing the Public Interests: Environmental Conservation ...................................................... 322
   2. Weighing the Public Interests: Settled Expectations ............................................................... 322
CONCLUSION ............................................................................ 324

INTRODUCTION

California is in desperate need of housing.\(^1\) To address the housing shortage, various cities and developers have turned to former and underused golf courses. For example, the Riverwalk Golf Club in San Diego, California is being transformed into a mixed-use development and is expected to offer 930 apartment homes by 2025.\(^2\) However, most attempts at transforming golf courses into housing are not as successful. A developer who wanted to build 443 residences on the Westridge Golf Club in La Habra, California, sued the city of La Habra based on allegations that the city unlawfully blocked the project.\(^3\) Likewise, projects to build thirty-nine homes in Orange, California, on the former Ridgeline golf course and discussions to build affordable housing on Willowick Golf Course in Santa Ana, California, have come to a complete halt.\(^4\) Plenty of barriers must be removed in order to enable the success of these types of development projects—projects that transform old or

---

unused golf courses into housing for a housing-starved population. One such barrier is conservation easements.⁵

Many golf courses are protected by conservation easements,⁶ which are a form of negative easement where landowners donate their land to a public body or private charitable organization by agreeing not to develop their land in a way that contradicts the terms of the easements.⁷ Typically, conservation easements intend to protect the environment by preserving historic areas, scenic areas, and open space.⁸ In order to achieve this goal, conservation easements exist in perpetuity⁹ and can only be terminated in extremely narrow circumstances—if at all.¹⁰ Golf courses are not worthy of such high-level protection. Not only are golf courses artificial nature, but they also actively harm the environment.¹¹ Additionally, golf is declining in popularity, and golf course use is dwindling—many golf courses are even shutting down.¹² Golf courses inherently require hundreds of acres of land, but with fewer and fewer people using golf courses, whether golf courses are worth the land they occupy is becoming increasingly unclear. Meanwhile, California faces a huge housing shortage,¹³ and Californians relinquish their dreams of homeownership, struggle endlessly to pay their rents, or are forced to live on the streets as a result.¹⁴ In order to encourage the productive use of land and address California’s housing shortage, conservation easements on golf courses should be terminable such that lesser-used golf courses can be developed into housing.

This Note argues that golf courses are undeserving of the continued protection of conservation easements and that the termination of golf course conservation easements should be made

---

⁵ For the purposes of this paper, conservation easements, open-space easements, and conservation servitudes all refer to the same land use restrictions. For the most part, I will be using the term conservation easements throughout this paper.
⁶ See infra Part III.A.
⁸ See id.
⁹ See CAL. GOV’T CODE § 51051(a) (West 2013); CAL. GOV’T CODE § 51070 (West 1977); CAL. GOV’T CODE § 51075(d) (West 1977); CAL. CIV. CODE § 815.2(b) (West 1979).
¹⁰ See CAL. GOV’T CODE § 51061 (West 1971); CAL. GOV’T CODE § 51093(a) (West 1974); THOMAS S. BARRETT & PUTNAM LIVERMORE, THE CONSERVATION EASEMENT IN CALIFORNIA 23 (Linda Gunnarson ed., 1983).
¹¹ See infra Part II.B.2.
¹² See infra Part II.A.
¹³ See infra Part I.
possible. The starting point of this argument is California’s housing shortage and its need for housing development. Moreover, the termination of golf course conservation easements is justified, if not necessitated, by the fact that golf courses serve an ever-shrinking population and, more importantly, do not serve a legitimate conservation easement purpose. In light of golf course conservation easements being an unproductive use of land, developing housing on golf courses is a convenient way to address California’s housing crisis. In order to make housing development on golf courses encumbered by conservation easements possible, golf course conservation easements must be terminated. However, conservation easements are extremely difficult, if not impossible, to terminate because conservation easements are perpetual—some actually perpetual, others constructively perpetual. This Note concludes by proposing steps that the California Legislature and the courts can take to enable termination of golf course conservation easements.

This Note consists of four parts. Following this introduction, Part I describes past and present housing situations in California. California’s housing history sheds light on the social environment that motivated the enactment of the laws that allow for the creation of conservation easements, as well as the stark difference between California’s past housing landscape and its current housing crisis. This idea demonstrates why California can no longer afford to protect artificial environments like golf courses with conservation easements. Part II argues that golf courses are an unproductive use of land—unworthy of conservation easement protection—and contends that replacing those golf courses with housing would be a productive use of land. Generally, golf courses are an unproductive use of land because golf is declining in popularity. Specifically, golf courses are undeserving of conservation easement protection because golf courses serve neither the statutory intent nor the environmental objective of conservation easements. On the other hand, developing housing on these golf courses is a productive use of land because it fulfills two public policy goals. Part III establishes that a major barrier to transforming golf courses encumbered by conservation easements into housing is the perpetuity feature of conservation easements and explains why perpetual conservation easements are harmful. Finally, Part IV proposes that the California Legislature and the

---

15 See infra Part II.A.
16 See infra Part II.B.
17 See infra Part III.A.
courts enable the termination of golf course conservation easements to make way for housing development.

I. HOUSING IN CALIFORNIA: PAST AND PRESENT

In order to understand why conservation easements should not continue to be used as a barrier to housing development, it is important to understand that the California Legislature did not contemplate the modern-day housing crisis. We must consider how Californians viewed housing during the 1960s and 1970s, which is when the California Legislature contemplated the legislation that forms the current conservation easement landscape.\(^{18}\) Contrasting these past views with the changes that occurred between the 1970s and the present as well as with the current housing crisis unveils the necessity of change in California’s land use priorities: California can no longer afford to prioritize open space over housing development.

From 1940 to 1960, California’s population jumped from 6.9 million to 15.7 million.\(^{19}\) As California’s population drastically increased, Californians watched their neighborhoods change before their eyes.\(^{20}\) This led to the antigrowth movement because Californians believed newcomers were increasing traffic, overwhelming infrastructure, and eliminating open space.\(^{21}\) By the mid-1960s, this antigrowth movement was in full swing,\(^{22}\) and one conservationist even argued that California could oppose growth by not building housing for newcomers.\(^{23}\) At the same time, the homeownership rate in America drastically increased by 1970, and in 1966, there were more Americans living in the suburbs than in the cities for the first time ever.\(^{24}\) As more and more American homeowners acquired and grew attached to open backyards, they increasingly feared that further housing development would take these backyards away.\(^{25}\) At the same time, the inflation that

\(^{18}\) See infra Part III.A.
\(^{21}\) Dougherty, supra note 19, at 69.
\(^{22}\) Id. at 79.
\(^{23}\) See Raymond F. Dasmann, The Destruction of California 190 (First Collier Books 1966) (1965) (“People will not come where there are no new jobs or new housing, or if they do come they will not stay.”).
\(^{24}\) Dougherty, supra note 19, at 81.
\(^{25}\) Id.
dominated the American economy beginning in the mid-1960s.\textsuperscript{26} While this devastated Americans trying to enter the housing market, it also made houses extremely valuable assets for existing homeowners.\textsuperscript{27} Consequently, American homeowners once again became increasingly opposed to nearby housing development—this time, because increasing the supply of housing decreased the value of their most valuable assets.\textsuperscript{28} By the end of the 1970s, real estate was a significant contributor to American household wealth, and those living in the suburbs more openly opposed housing development.\textsuperscript{29} The drastic increases in California’s population and the growing notion that homeowners had to staunchly protect their homes fostered a desire to stop California from changing.\textsuperscript{30} This was the state of housing in California when the California Legislature enacted conservation easement legislation to prevent development in the state.\textsuperscript{31}

Despite the best efforts of California homeowners and the California Legislature, the state’s population, and, thus, its housing needs, have changed. Between 1960 and 1970, California’s population increased from 15.7 million to 20 million.\textsuperscript{32} As of 2020, California’s population was 39.5 million.\textsuperscript{33} Meanwhile, housing development continually decreased.\textsuperscript{34} In the 1970s, an average of 215,585 building permits were issued every year for housing construction in California.\textsuperscript{35} For the period from 2016 to 2021, that average fell to 110,474 building permits per year.\textsuperscript{36}

\begin{footnotes}
\item[27] DOUGHERTY, supra note 19, at 85.
\item[28] Id.
\item[29] Id.
\item[30] See id.
\item[31] See id.
\item[32] See Julia D. Mahoney, Perpetual Restrictions on Land and the Problem of the Future, 88 VA. L. REV. 739, 751 (2002) (“Many landowners are delighted by the thought that the land to which they have formed emotional attachments will remain as they know and love it [due to the imposition of conservation easements].”).
\item[34] See id.
\item[36] See id.
\end{footnotes}
California’s housing supply has not changed much since 2008, despite California’s population increasing by 6.1% between 2010 and 2020. To be more specific, only 24.7 new units were built for every 100 new residents between 2007 and 2017. In part, this paradox is a result of existing homeowners resisting new development in order to keep their own property values high. However, California’s strict and numerous zoning laws also contribute to this paradox because they make housing development quite costly for developers, and developers often pass those costs down to homeowners. Overall, this paradox has made Californian housing even less affordable and increased California’s unhoused population.

California is the most expensive state in which to buy a home. Californians cannot even avoid this expense by renting homes because California’s rental rates are also constantly growing. Nearly half of Californian households cannot afford housing in their local markets. Unsurprisingly, lower-income Californians, who have to spend a greater portion of their incomes on housing, are harmed the most by California’s housing crisis. In fact, low-income Californians account for 38% of California’s population, and almost none of them can afford housing in California. However, the housing crisis has also made homeownership extraordinarily challenging for middle-income Californians.

The cost of buying or renting a home is not the only challenge that the limited supply of housing imposes on Californians. California has greater housing costs—i.e., mortgage payments, property taxes, and maintenance—for both homeowners and

39 See U.S. Census Bureau, supra note 33, at 1.
40 See Perry et al., supra note 38, at 18. This is a little less than half of the 43.1 new units per resident built nationwide. See id. To frame this shortage another way, between 2011 and 2016, California added 135,000 more households than housing units. See id. at 26.
41 See id. at 18.
42 See id. at 18–19.
43 See id. at 16. To be precise, California has approximately 130,000 unhoused people, which is the largest population of unhoused people in America by a large margin. See Dougherty, supra note 19, at 152.
44 See Perry et al., supra note 38, at 26.
45 See id.
46 See Jonathan Woetzel et al., A Tool Kit to Close California’s Housing Gap: 3.5 Million Homes by 2025 4 (McKinsey Global Institute eds., 2016).
47 See Perry et al., supra note 38, at 26.
48 See Woetzel et al., supra note 46, at 5.
49 See Perry et al., supra note 38, at 26.
renters, in comparison to other states. Specifically, California “has the highest percentage of house-burdened households among homeowners”—exceeding the American average by over 10%. A household is house-burdened if housing costs constitute at least 30% of the household income. Similarly, California has the highest percentage of rent-burdened households in America. Even worse, almost 70% of low-income households in California are extremely cost burdened, which means they spend at least 50% of the household income on housing.

Clearly, California needs to increase its housing supply in order to make both homeownership and renting affordable. More precisely, the McKinsey Global Institute estimated that, as of October 2016, California had two million fewer housing units than it needed, and this estimate was considered conservative.

II. GOLF COURSES CONSERVATION EASEMENTS: AN UNPRODUCTIVE USE OF LAND

One way California can increase its housing supply is by terminating golf course conservation easements. Not only is golf on the decline—resulting in the decreased use of the exorbitant amounts of land golf courses command, but golf courses do not serve any legitimate conservation easement purposes. Terminating golf course conservation easements would increase the opportunities for developers to build housing that California desperately needs, and, unlike golf courses, housing development on former golf courses would actually help the environment.

A. The Declining Popularity of Golf

For the better part of the 21st century, golf has seen a drastic decline in popularity in America. According to the National Recreation and Park Association, the population of golfers

---

50 See id. at 7.
51 See id.
52 See id.
53 See id. In 2016, the median gross rent in California was 40.2% higher than the American average. See id. at 10. The median single-family home price in California was 113.3% higher than the American average. See id.
54 See WOETZEL ET AL., supra note 46, at 5.
55 See PERRY ET AL., supra note 38, at 27.
56 See WOETZEL ET AL., supra note 46, at 2.
57 Id. at 2. See also PERRY ET AL., supra note 38, at 19 n.20.
decreased by twenty-two percent between 2003 and 2018.\textsuperscript{59} Granted, golf saw an uptick in popularity in the summer and fall of 2020 during the pandemic because it is an outdoor sport that allowed players to socially distance.\textsuperscript{60} However, as early as the summer of 2021, the golf community has seen indications that this trend has already started to fade.\textsuperscript{61}

Golf’s decline in popularity is primarily a function of golf being too expensive and too time consuming.\textsuperscript{62} Golf has been an expensive sport for some time now due to the equipment required to play the game and the cost of tee times,\textsuperscript{63} and it has only gotten more expensive over time.\textsuperscript{64} Beyond cost, golf is too time consuming to be compatible with the busy lifestyles of most Americans.\textsuperscript{65} Most Americans work over forty hours per week, leaving them with few opportunities to spend several hours playing a round of golf.\textsuperscript{66}

\textsuperscript{59} Id.
\textsuperscript{65} See Fitzpatrick, supra note 62.
\textsuperscript{66} See \textit{id}.
Golf courses also take up a lot of space; the average 18-hole golf course requires 150 acres of land.67 In total, golf courses occupy up to two million acres of land in America68 and seventy thousand acres of land in Southern California alone.69 While that might be tolerable for golf courses that serve a lot of players, there are plenty of golf courses that are not seeing many players anymore.70 Those golf courses simply are no longer a good use of valuable space—especially in light of other, more pressing land uses. Yet, they are still offered the protection of conservation easements.71

B. Golf Courses Do Not Serve a Legitimate Conservation Easement Purpose

1. Golf Courses Do Not Serve the Statutory Intent of Conservation Easements

According to the Open-Space Easement Act of 1974 (the “1974 Act”), open-space land refers to land that “is essentially unimproved and devoted to an open-space use as defined in Section 65560 of the Government Code.”72 Under Section 65560(h), the following activities are considered open-space uses: preservation of natural resources; managed production of resources; outdoor recreation; public health and safety; support of military installations; and protection of Native American places, features, and objects.73

Because golf courses do not produce resources; require regulation due to public health and safety; support military installations; or protect Native American places, features, and objects, golf courses only have potential to serve an open-space use


70 See Crompton, supra note 58 (noting that “[i]n a typical year, approximately 200 courses fail”); Baughman, supra note 64 (declaring that as of 2017, “more than 800 courses across America have closed in a decade”).


72 CAL. GOV’T CODE § 51075(a) (West 1977). Section 51075(a) refers to Section 65560 for a definition of open-space use. Id.

73 CAL. GOV’T CODE § 65560(b)(1)–(6) (West 2018).
as preservation of natural resources or outdoor recreation under the 1974 Act. However, closer examination of the meaning of open space for the preservation of natural resources and open space for outdoor recreation raises the question of whether the California Legislature intended for open-space land to encompass golf courses as we understand them today. Section 65560(h)(1) describes “open space for the preservation of natural resources” as including “areas required for the preservation of plant and animal life, including habitat for fish and wildlife species; areas required for ecologic and other scientific study purposes; rivers, streams, bays, and estuaries; and coastal beaches, lakeshores, banks of rivers and streams, and watershed lands.” 74 Golf courses do not preserve plant and animal life because in order to construct golf courses, natural habitats, and the plants and animals that rely on them, must be “disfigured and destroyed to create highly organized, artificially watered and unarguably fake nature.” 75 For the same reason, golf courses are not useful for ecologic and other scientific study. The artificial character of golf courses means that golf courses rarely contain natural bodies of water like rivers, streams, bays, and estuaries. 76 This also means that golf courses rarely contain coastal beaches, lakeshores, banks of rivers and streams, and watershed lands. 77 Thus, golf courses cannot serve as open space for the preservation of natural resources. Moreover, Section 65560(h)(3) clarifies:

Open space for outdoor recreation[] includ[es] . . . areas of outstanding scenic, historic, and cultural value; areas particularly suited for park and recreation purposes, including access to lakeshores, beaches, and rivers and streams; and areas that serve as links between major recreation and open-space reservations, including utility easements, banks of rivers and streams, trails, and scenic highway corridors. 78

Golf courses definitely do not have historic or cultural value. Additionally, while golf courses do serve a recreational purpose, because the list of park and recreation purposes provided in Section 65560(h)(3) exclusively contains bodies of water, it is doubtful that the California Legislature intended for Section 65560(h)(3) to include golf-related recreation because most golf courses do not provide access to lakeshores, beaches, and rivers

74 Id.
75 See Baughman, supra note 64.
77 See id.
78 CAL. GOV'T CODE § 65560(h)(3) (West 2018).
and streams.\(^{79}\) Regardless of whether golf courses serve any scenic value or serve as links between major recreation and open-space reservations, at the end of the day, the 1974 Act requires that open-space land be “essentially unimproved,”\(^{80}\) and golf course construction requires drastically changing the land on which golf courses are built.\(^{81}\)

2. Golf Courses Do Not Serve the Environmental Objective of Conservation Easements

Even accepting the assumption that conservation easements are a productive use of land would be of no help to preserving golf courses because it would be based on the argument that conservation easements promote environmentalism.\(^{82}\) Yet, many conservation easements protect golf courses even though golf course construction and maintenance harm the environment in a multitude of ways. In the process of clearing land for golf courses, golf course developers remove vast amounts of natural vegetation and habitats from the land, which “ravage[s] entire ecosystems” and destroys biodiversity.\(^{83}\) Specifically, clearing the land requires excavation and soil movement, which alter natural habitats.\(^{84}\) After destroying native vegetation and driving out animals that inhabited the land, golf course developers fill the land with non-native grasses and decorative plants.\(^{85}\)

---

\(^{79}\) The canon of statutory construction known as *noscitur a sociis* dictates that a group of words should take on similar meanings in order to avoid inadvertently widening the scope of statutes. See NORMAN J. SINGER & SHAMIE SINGER, 2A SUTHERLAND STATUTES AND STATUTORY CONSTRUCTION § 47:16 (7th ed. 2014).

\(^{80}\) CAL. GOV'T CODE § 51075(a) (West 1977).

\(^{81}\) See Baughman, supra note 64; see also infra notes 85–87 and accompanying text.


\(^{85}\) See BEACHAPEDIA, supra note 83.
To make matters worse, the negative environmental impacts of golf courses continue beyond the construction of such golf courses. Golf courses require excessive amounts of water to keep acres of grass green.86 Oftentimes, the local water supply cannot afford to meet the demands of golf courses.87 For example, in California, a notorious drought state,88 thirty-seven million gallons of water are used on a daily basis to water golf courses in Coachella Valley alone.89 Land that draws so intensely upon such a limited resource surely cannot be considered natural or environmentally friendly. Additionally, in order to maintain the strength of the grass on golf courses, golf course managers must feed nutrients into the grass through unnatural fertilizers.90 They also spread pesticides all over golf courses in order to protect the grass and keep pests away from golfers.91 There are usually regulations on the amount of pesticides golf course managers can use, but “a vast majority of golf course managers ignore the international and local regulations in terms of pesticides and [the] workers who handle them do not know the rules.”92 In fact, more pesticides are applied to golf courses than farmlands.93 The harm done by fertilizers and pesticides expands beyond the golf courses. These chemicals get into the irrigation systems and contaminate groundwater aquifers, surface water bodies, and the ocean.94

From destroying local biodiversity, to using an exorbitant amount of California’s limited water supply, to polluting bodies of water with pesticides and fertilizers, golf courses cause an abundance of environmental harm and do not serve a legitimate conservation easement purpose. Therefore, even assuming conservation easements are a productive use of land, golf course conservation easements are not a productive use of land.

86 Rosenberg, supra note 83.
87 See BEACHAPEDIA, supra note 83.
89 Baughman, supra note 64. California is not the only drought state where this happens. In Phoenix, Arizona, golf courses use more than eighty million gallons of water per day. Id. Similarly, in Salt Lake City, Utah, thirty golf courses consume approximately nine million gallons of water per day. Rosenberg, supra note 83.
90 See Guzmán & Fernández, supra note 84, at 418.
91 See id.; see also Rosenberg, supra note 83.
92 Guzmán & Fernández, supra note 84, at 418.
93 See id. at 419.
94 BEACHAPEDIA, supra note 83; see also Baughman, supra note 64; Rosenberg, supra note 83; Guzmán & Fernández, supra note 84, at 419.
C. Replacing Golf Courses with Housing Is a Productive Use of Land

Terminating conservation easements on golf courses to enable housing development removes unproductive land use and replaces it with productive land uses. For obvious reasons, replacing golf courses with housing fulfills the public policy goal of addressing the housing crisis, but developing housing on golf courses also fulfills environmental policy goals. This is because transforming golf courses into housing enables development near the already-existing developments.95

California needs housing, so homes must be built one way or another.96 Housing can either be built in areas that are already developed—such as suburban neighborhoods with increasingly vacant golf courses—or in areas that are currently untouched by societal developments. Building homes in areas that are already developed is better for the environment because it preserves natural landscapes, reduces greenhouse gases, and decreases pollution from water runoff.97 First, if housing is not constructed near existing developments, by default, it must be constructed in distant, undeveloped areas. This sprawl inherently requires the destruction of untouched ecosystems. Second, when housing is built away from jobs, services, and other homes, residents of those homes are generally forced to drive further to get to work, run errands, and meet up with friends and family.98 That is, when homes are built away from existing developments, residents of those homes have longer commutes, which generally increases greenhouse gas emissions.99 On the other hand, building housing near existing developments reduces greenhouse gas emissions.100 Finally, developing in previously untouched areas creates

95 See infra text accompanying notes 92–98.
96 See U.S. Env’t Prot. Agency, EPA 231-R-06-001, Protecting Water Resources with Higher-Density Development (2006) (“[T]he choice is not whether to grow by one house or eight but is instead where and how to accommodate the eight houses.”).
100 See Decker, supra note 98.
impervious surfaces and compacted soils that filter less water. This increases surface runoff and decreases groundwater infiltration, which results in the pollution of streams, rivers, lakes, and beaches. Transforming golf courses into housing minimizes the need for environmentally harmful sprawl.

Overall, not only do golf courses fail to fulfill an environmental preservation purpose, but they actively harm the environment. As a result, golf courses are undeserving of protection by conservation easements because they do not pursue the public policy goals of environmentalism that conservation easements purport to accomplish. Meanwhile, building housing in developed areas is an important step in addressing California’s housing crisis and has a positive impact on the environment. Clearly, terminating golf course conservation easements to enable the transformation of golf courses into housing is an efficient way to address California’s housing crisis and promote environmentalism. However, one enormous barrier stands in the way: the perpetuity feature of conservation easements.

III. THE PERPETUITY FEATURE OF CONSERVATION EASEMENTS: THE CONS OUTWEIGH THE PROS

One of the most controversial components of conservation easements is the perpetuity feature. Yet, all three of California’s conservation easement acts are actually or constructively perpetual. Although the perpetuity feature of conservation easements has its perks, the harm caused by this feature drastically outweighs those perks. Namely, in addition to not

---

102 U.S. ENV’T PROT. AGENCY, supra note 96, at 4.
103 Id.
105 See infra Part III.A.
106 See infra Part III.B.
actually being an effective way to conserve nature, perpetual conservation easements bind future generations to outdated scientific knowledge and cultural values.\textsuperscript{107}

A. All Conservation Easements Are Actually or Constructively Perpetual

When discussing conservation easements in California, there are three key pieces of legislation to recognize: the Conservation Easements Act of 1979 (the “1979 Act”),\textsuperscript{108} the Open-Space Easement Act of 1974 (the “1974 Act”),\textsuperscript{109} and the Open-Space Easement Act of 1969 (the “1969 Act”).\textsuperscript{110} While there are differences between the conservation easements created under the 1979, 1974, and 1969 Acts, landowners, for the most part, can pursue land conservation under any of these three acts.\textsuperscript{111} Terminating a conservation easement, on the other hand, requires significantly more thought.

To start, there are currently no means by which a conservation easement granted under the 1979 Act can end.\textsuperscript{112} Conservation easements under the 1979 Act \textit{must} be granted in perpetuity,\textsuperscript{113} and the 1979 Act has no termination provision.\textsuperscript{114} Therefore, there is currently no statutory method for terminating a 1979 Act conservation easement. Accordingly, conservation easements granted under the 1979 Act are actually perpetual.

While conservation easements under the 1969 and 1974 Acts can be perpetual or for a term of years and can be nonrenewed or terminated via statutorily prescribed methods, the requirements for nonrenewal and termination render the conservation easements under the 1969 and 1974 Acts constructively perpetual.\textsuperscript{115} Conservation easements granted under the 1974 Act can be approved for a minimum term of ten years, but can also exist in perpetuity.\textsuperscript{116} A perpetual conservation easement can only be terminated by abandonment under Section 51093.\textsuperscript{117} Pursuant to Section 51093, a city or county may only approve of

\textsuperscript{107} See infra Part III.C.
\textsuperscript{108} CAL. CIV. CODE §§ 815–816 (West 1979).
\textsuperscript{109} CAL. GOV. CODE §§ 51070–51097 (West 1974).
\textsuperscript{110} CAL. GOV. CODE §§ 51050-51065 (West 1969).
\textsuperscript{111} See infra notes 108–121 and accompanying text.
\textsuperscript{112} See infra notes 115–116 and accompanying text.
\textsuperscript{113} CAL. CIV. CODE § 815.2 (West 1979).
\textsuperscript{114} See BARRETT & LIVERMORE, supra note 10, at 32.
\textsuperscript{115} See infra notes 110–121 and accompanying text.
\textsuperscript{116} CAL. GOV’T CODE § 51070 (West 1977); see also CAL. GOV’T CODE § 51075(d) (West 1977); CAL. GOV’T CODE § 51081 (West 1975).
\textsuperscript{117} See BARRETT & LIVERMORE, supra note 10, at 24.
Challenging Subpar Servitudes

an abandonment if it finds that the conservation easement no longer serves any Section 51084 public interest, the abandonment is consistent with the purpose of the 1974 Act, the abandonment is consistent with the local general plan, and the abandonment is necessary to avoid a substantial financial hardship to the landowner due to involuntary factors unique to the landowner. Not only are the latter three requirements vague and easy to oppose, but Section 51084 also covers a broad range of public interests. As a result, it is unlikely that land that once served one of these public purposes no longer serves any Section 51084 public purpose; thus, abandoning a 1974 Act conservation easement is unlikely to be successful. Furthermore, a conservation easement for a term of years can be terminated by abandonment under Section 51093 or by nonrenewal. However, if nonrenewal is initiated by a nonprofit organization, the nonrenewal must be approved by the city or county according to the standards provided in Section 51093, which makes nonrenewal just as impossible as abandonment. Clearly, conservation easements granted under the 1974 Act are constructively perpetual.

The termination process is extremely similar for conservation easements granted under the 1969 Act. According to Section 51053 of the 1969 Act, a conservation easement must last for at least twenty years, but can also exist in perpetuity. This conservation easement may only be abandoned if the city or county first finds that the land no longer serves a public purpose listed in Section 51056(b). However, Section 51056(b) has a fairly broad list of public purposes, so it is unlikely that land that once

118 A city or county must find that preservation is in the best interest of the city or county. CAL. GOV’T CODE § 51084(b) (West 2013). Moreover, at least one of the following must be true: the land is essentially unimproved and has value as scenery, a watershed, or as a wildlife preserve in its natural state; the land will add to the amenities of living in neighboring developed areas or will help preserve rural character; the land lies in an area that “should remain rural in character” and keeping the land as open space will help maintain that character; the land will prevent floods or has value as a watershed; the land rests in “an established scenic highway corridor”; the land is a wildlife preserve or sanctuary; and the open space will serve the purposes of the 1974 Act or Section 8 of Article XIII of the California Constitution. Id.
119 CAL. GOV’T CODE § 51093(a) (West 1974).
120 CAL. GOV’T CODE § 51093(a) (West 1974).
121 CAL. GOV’T CODE § 51090 (West 1977).
122 CAL. GOV’T CODE § 51090 (West 1977).
123 CAL. GOV’T CODE § 51053 (West 2012).
124 CAL. GOV’T CODE § 51051(a) (West 2013).
125 CAL. GOV’T CODE § 51061 (West 1971).
126 The Section 51056(b) public purposes are as follows:
served one of these public purposes no longer serves any Section 51056(b) public purpose. Consequently, abandoning a 1969 Act conservation easement is unlikely to be successful, making a 1969 Act conservation easement constructively perpetual.

Overall, between the 1969 Act, the 1974 Act, and the 1979 Act, there are plenty of ways to impose permanent conservation easements on a wide variety of lands. However, there is no reasonable way to terminate the conservation easements and free those lands should needs or interests change.

B. The Perpetuity Feature of Conservation Easements Serves Only a Few Purposes

Facially, it makes sense that conservation easements are perpetual. The purpose of conservation easements is to preserve the environment by restricting development on the land. If land development is only restricted for a few years, or even a few decades, then eventually the land will be developed. The idea is that once the land is developed, the environment can no longer be preserved. Thus, the only way to preserve the environment is by ensuring that

---

[T]he preservation of the land as open space is in the best interest of the state, county, or city and is important to the public for the enjoyment of scenic beauty, for the use of natural resources, for recreation, or for the production of food or fiber and specifically because one or more of the following reasons exist:

1. It is likely that at some time the public may acquire the land for a park or other public use.
2. The land is unimproved and has scenic value to the public as viewed from a public highway or from public or private buildings.
3. The retention of the land as open space will add to the amenities of living in adjoining or neighboring urbanized areas.
4. The land lies in an area which in the public interest should remain rural in character and the retention of the land as open space will help preserve the rural character of the area.
5. It is in the public interest that the land remain in its natural state, including the trees and other natural growth, as a means of preventing floods or because of its value as watershed.
6. The land lies within an established scenic highway corridor.
7. The land is valuable to the public as a wildlife preserve or sanctuary and the instrument contains appropriate covenants to that end.
8. The public interest will otherwise be served in a manner recited in the resolution and consistent with the purposes of this subdivision and Article XXVIII of the Constitution of the State of California.

CAL. GOVT CODE § 51056(b) (West 1969).

127 See, e.g., Schwing, supra note 104, at 218 (describing termination of conservation easements as “extremely difficult or impossible”).

Challenging Subpar Servitudes

the land is never developed. The perpetuity feature makes conservation easements a stronger land preservation mechanism than regulations.\textsuperscript{129} Conservation easements create an unwavering right in the preservation of the land, and this right is less likely to falter in the face of changing political and economic needs and interests than regulations, which are always subject to change.\textsuperscript{130}

Beyond environmental preservation purposes, the tax considerations favor the imposition of the perpetuity feature of conservation easements. When landowners subject their land to conservation easements, the I.R.S. views it as a charitable contribution and allows the landowners to take tax deductions.\textsuperscript{131} This tax deduction incentivizes landowners to burden their land for the sake of environmental preservation.\textsuperscript{132} However, if terminating conservation easements were convenient, landowners could exploit the tax deductions and then swiftly unburden their land—enabling landowners to take advantage of a reward they have done nothing to earn.\textsuperscript{133}

Finally, the perpetuity feature actually propels the conservation easement movement forward by appealing to nostalgia, as landowners tend to have personal attachments to their land and do not want to see their land change.\textsuperscript{134} Through conservation easements, landowners can attempt to stop time by ensuring that, even when they are long gone, their land will continue to resemble the picture in their memory.\textsuperscript{135}

C. Perpetual Conservation Easements are Harmful Because They are Ineffective and Bind Future Generations to Outdated Science and Cultural Values

When past and present landowners grant conservation easements, they make decisions that are extraordinarily difficult for future landowners to reverse. Understandably, many scholars have

\begin{itemize}
\item \textsuperscript{130} See id.
\item \textsuperscript{131} See I.R.C. § 170(a) (allowing deductions for charitable contributions); I.R.C. § 170(f)(3)(B)(ii) (asserting that taxpayers are not denied deductions for qualified conservation contributions).
\item \textsuperscript{132} See Richard J. Roddewig, \textit{Conservation Easements & Their Critics: Is Perpetuity Truly Forever . . . and Should It Be?}, 52 UIC J. Marshall L. Rev. 677, 682 (2019); id. at 682 n.8.
\item \textsuperscript{133} See Nancy A. McLaughlin, \textit{Conservation Easements: Perpetuity and Beyond}, 34 Ecology L. Q. 673, 675–76 (2007). In fact, land trusts trying to acquire conservation easements use this as a selling point with landowners. \textit{Id.} at 676; see also Mahoney, supra note 32, at 750–51.
\item \textsuperscript{134} See Nancy A. McLaughlin, \textit{Conservation Easements: Perpetuity and Beyond}, 34 Ecology L. Q. 673, 675–76 (2007). In fact, land trusts trying to acquire conservation easements use this as a selling point with landowners. \textit{Id.} at 676; see also Mahoney, supra note 32, at 750–51.
\item \textsuperscript{135} See Mahoney, supra note 32, at 750–51.
\end{itemize}
qualms with this. Through conservation easements, past and present landowners exercise control over future landowners and have the ability to impose outdated land use ideas on a community whose needs and interests have evolved and can no longer be served by those land use ideas. This leaves future generations stuck with restrictions that do not promote modern values or incorporate advances in ecological sciences. Not only do past and present landowners violate the autonomy of future landowners when they grant perpetual conservation easements in an attempt to save the environment, they are improperly assuming that they hold all the answers to future problems. It is naive to assume that past and present generations know more than future generations. Yet, humans regularly assume that the information they have constitutes “enduring truths instead of contingent hypotheses.” They tend to overestimate their competence and forget that future generations inevitably change the plans of past and present generations because, undoubtedly, future generations will have a greater understanding of how effectively conservation easements actually contribute to land conservation.

In a 2002 article entitled “Perpetual Restrictions on Land and the Problem of the Future,” Julia D. Mahoney provides several arguments supporting the idea that past and present generations do not know enough to make permanent decisions for future generations—three of which support the argument that conservation easements on golf courses no longer serve public policy goals. First, Mahoney debunks the argument that the best way to preserve nature is to not touch it at all. This old argument derived from the homeostasis model, which maintained that living organisms and their habitats persisted by resisting change. Clearly, conservation easements and their perpetual nature are backed by the homeostasis model.
In fact, the homeostasis model formed the basis of environmentalism in the mid-1960s—just before the California legislature enacted the 1969, 1974, and 1979 Acts. However, ecological scientists have abandoned the homeostasis model for the belief that nature is in a constant state of flux. Because nature is in a constant state of flux, it is impossible for humans to preserve the earth by simply maintaining it as is. Thus, perpetual conservation easements, like the homeostasis model on which they are based, are outdated. The people involved in preserving land through conservation easements incorrectly believe that they are helping the earth by preventing development on these protected lands, but in reality, “the eternal prohibition of residential subdivisions, commercial activity, and other ‘development’ may turn out to be foolish.”

Second, Mahoney explains that the perpetuity feature is harmful because future generations will inevitably better understand ecosystems and environmental preservation than past and present generations. With respect to golf course conservation easements, the advancement in understanding of what Mahoney describes is evident. To start, improved knowledge about how ecosystems work has changed views on golf courses as natural environments. As recently as fifty years ago, even golf course designers did not know how golf courses impacted the environment. However, it is now known that the construction and maintenance of golf courses actually destroy biodiversity and pollute irrigation systems, among other harms. Furthermore, in the 1970s, Dr. Paul R. Ehrlich’s *The Population Bomb* initiated an anti-growth movement on the basis that “[t]oo many people, packed into too-tight spaces, [took] too much from the earth.” This anti-growth movement partially manifested in opposition to development, which inevitably made big open spaces like golf courses very desirable. Over time,

---

147 See Dougherty, *supra* note 19, at 80.
148 Mahoney, *supra* note 32, at 754.
149 Id.
150 Id. at 757.
151 Id.
152 Yi, *supra* note 69.
153 See *supra* Part II.B.2.
Dr. Ehrlich’s view has become rather fringe, and it is now known that the overpopulation Dr. Ehrlich described can be attributed to economics and sociology rather than science.156 Because it is now known that golf courses do not preserve ecosystems and the anti-growth movement does not preserve the environment, the reasons for protecting golf courses with conservation easements are moot. Yet, due to the current legislation surrounding conservation easements, land uses cannot be improved in response to this change. Even worse? In all likelihood, future generations will learn even more about environmentalism and wish to change the approach that past and present generations have taken to preservation—including the use of conservation easements.157 Sadly, with the existing rules surrounding conservation easements, future generations will find changing the approach extraordinarily onerous, if not impossible.

Third, Mahoney argues that perpetual conservation easements are detrimental because cultural values change from generation to generation and further suggests that those changing cultural values can effect change in future land use regulations.158 For example, in 1958, Pat Brown ran for governor of California with a campaign focused on growing California’s population, and Californians elected Brown as governor by a margin of a million votes.159 Through their voting, Californians indicated that they favored growth.160 However, by the mid-1960s, Californians shifted their preference, and anti-growth sentiment took over local politics.161 Californians have changed their preferences before, so another shift is entirely possible. As the popularity of golf declines and California’s need for housing increases, the cultural values of Californians may shift away from protecting acres of golf courses and instead favor developing land to make housing possible for their fellow Californians. On a more general level, perhaps future generations will not oppose development with as much voracity as the current generation opposes development.162 Still, with perpetual conservation easements in the way, that will not matter because future generations will not have the liberty to align their land uses to their cultural values. Notably, the Open-Space Easement Act of 1974 itself presumes that public policy interests

156 Mann, supra note 154.
157 See Mahoney, supra note 32, at 757.
158 Id. at 759.
159 DOUGHERTY, supra note 19, at 69.
160 Id.
161 Id. at 79.
162 Mahoney, supra note 32, at 762.
will change. Section 51084 states that a conservation cannot be granted unless there is a public interest, but Section 51093(a) says that a conservation cannot be terminated unless there is no Section 51084 public interest. Section 51093(a) inherently assumes that the Section 51084 public interest can change because if it is impossible for the Section 51084 public interest to change, then it would be impossible for a conservation easement to be terminated under Section 51093(a). In which case, Section 51093(a) is null, and the canons of statutory construction—namely, the rule against surplusage—require that meaning be given to every word of a statute. Clearly, changes in cultural values and public interests are inevitable and expected. Yet, the past and present generations continue to insist on making conservation easements perpetual. Unfortunately, by doing so, the past and present generations lock in the land and shut the door on future generations’ ability to evolve.

Throughout history, society has experienced shifts in economics, population, technology, and values that modify land use desirability. Even Restatement (Third) of Property: Servitudes Section 7.11, which strongly favors maintaining the perpetuity of conservation easements, recognizes that “it is inevitable that, over time, changes will take place that will make it impracticable or impossible for some conservation servitudes to accomplish the purpose they were designed to serve.” Overall, imposing on future generations restrictions that are likely to become outdated does not seem to work in favor of those future generations. If that’s the case, then what is the point? The entire justification for conservation easements is to preserve the land for future generations. If future generations cannot benefit, then what is the real justification? In her article, Mahoney speculates that “the real beneficiaries [of conservation easements] are members of the present generation.” That is, these beneficiaries are past...
and present generations who believe that leaving land undeveloped—or, in the case of golf courses, leaving land developed, but according to the preferences of modern people—is the best use of the land. However, this benefit comes at a great expense to future generations, and feeding into the fleeting desires of past and present generations is certainly not worth the suffering of future generations.

IV. PROPOSAL

While some conservation easements actually promote environmental conservation by safeguarding California’s coasts, forests, and historic areas, other conservation easements merely prevent acres of pesticide-soaked grass from being transformed into productive uses of land that benefit the community. Where conservation easements protect golf courses, conservation easements inhibit housing development and the environmental benefits associated with building housing near existing developments. Plain and simple: golf course conservation easements obstruct the productive use of land—especially because the number of people using golf courses is dwindling while the number of people needing housing is rising. Clearly, terminating golf course conservation easements can help California productively use its land, but the perpetuity feature of conservation easements obstructs this productivity. In order to grapple with the obstacle imposed by the perpetuity features, I present a two-part proposition. First, the California Legislature should disallow new grants of conservation easements on golf courses and enable termination of existing golf course conservation easements. Second, the courts should provide a means by which golf course conservation easements can be challenged and terminated.

A. The California Legislature Should Amend the 1979, 1974, and 1969 Acts

The California Legislature can disallow golf course conservation easements by amending the 1979, 1974, and 1969 Acts. The California Legislature should add a provision prohibiting conservation easements on land used as golf courses to each Act. Additionally, the 1979 and 1969 Acts define the type of land that qualifies for conservation easements very broadly, so the California Legislature should amend the 1979 and 1969 Acts to explicitly exclude land used as golf courses from conservation easements.

Challenging Subpar Servitudes

To remove existing barriers and enable the productive use of land, the California Legislature needs to create avenues through which existing golf course conservation easements can be terminated. This task can be accomplished by amending the 1979, 1974, and 1969 Acts to include provisions that explicitly allow conservation easements on golf courses to be terminated or abandoned. Alternatively, for the 1969 and 1974 Acts, the California Legislature can amend the abandonment provisions—Sections 51061 \(^{172}\) and 51093(a) \(^{173}\) respectively—to explicitly allow for the abandonment of conservation easements on golf courses.

B. The Courts Should Apply a Balancing Test to Terminate Conservation Easements

Another potential means for removing the barrier that golf course conservation easements pose to the productive use of land is a court-enforced equitable mechanism by which golf course conservation easements can be terminated. The California Supreme Court has held that equitable servitudes that “impose[] burdens on the affected land that are so disproportionate to the restriction’s beneficial effects” will not be enforced.\(^{174}\) This holding should apply to conservation easements and equitable servitudes\(^{175}\) alike because conservation easements, a type of negative easement,\(^{176}\) are more akin to covenants than easements and, thus, are typically analyzed as covenants.\(^{177}\)

Consequently, instead of enabling conservation easements to remain actually or constructively perpetual, courts should terminate conservation easements where the burdens that

\(^{172}\) CAL. GOV’T CODE § 51061 (West 1971).
\(^{173}\) CAL. GOV’T CODE § 51093(a) (West 1974).
\(^{175}\) In other words, an equitable servitude is a covenant enforceable in equity. Dukeminier, supra note 7, at 761.
\(^{177}\) Gerald Korngold, Private Land Use Arrangements: Easements, Real Covenants, and Equitable Servitudes 7 (1990). In fact, the Restatement (Third) of Property: Servitudes Section 1.2 explicitly declares that negative easements are indistinguishable from restrictive covenants and should be treated as restrictive covenants. RESTATEMENT (THIRD) OF PROP.: SERVITUDES § 1.2 cmt. b (AM. L. INST. 2000).
In the application of this balancing test, two interests will commonly oppose the productive use of land: environmental conservation and settled expectations.

1. Weighing the Public Interests: Environmental Conservation

As previously established, the public interest supporting the grant of conservation easements is environmental preservation. Because golf courses are an environmental harm, the public interest in golf course conservation easements will easily be outweighed by the public interest in land uses such as housing development. Meanwhile, conservation easements that actually preserve nature will be considered to serve a strong public interest and be continued. This balancing test not only encourages the productive use of land, but it also accommodates our evolving understanding of what constitutes good environmentalism because the balancing test enables people to challenge conservation easements and request that courts periodically reevaluate whether the conservation easements actually promote the public policy of environmental preservation.

2. Weighing the Public Interests: Settled Expectations

Although conservation easements are not granted with the intent of protecting settled expectations, many people have developed settled expectations around golf course conservation easements. In the last few decades of the twentieth century, golf course neighborhoods began to gain traction, but most people are not buying homes in golf course neighborhoods out of a love for golf. Rather, they enjoy seeing open, green space from their...
homes\textsuperscript{182} and reveling in the accompanying exclusivity.\textsuperscript{183} Homeowners in golf course neighborhoods get to look out their windows and see “gently rolling greens, clusters of mature trees, ponds, lakes and fountains, as well as an occasional wildlife sighting”\textsuperscript{184} instead of someone else’s house.\textsuperscript{185} Not only does the golf course increase these homeowners’ enjoyment of their homes, but it also increases the values of their homes.\textsuperscript{186}

In all likelihood, homeowners in golf course neighborhoods chose to purchase their homes because of the green space and exclusivity provided by the golf courses, and their willingness to pay the price tag on their homes was at least partially motivated by the existence of these benefits. Where golf courses are safeguarded by conservation easements, these homeowners have the expectation that they will continue to reap the golf-course-related benefits for which they paid. Thus, these homeowners—third parties to the conservation easements—have settled expectations that are dependent on the conservation easements. If those conservation easements are terminated and developers build housing on the golf courses, that would upset the settled expectations of these homeowners. Luckily for these homeowners, courts have demonstrated a willingness to defend the expectations that persuade homeowners to purchase homes.\textsuperscript{187}

Despite the courts’ willingness to defend settled expectations, there is still a strong argument to be made for terminating conservation easements in the interest of housing development. To start, it is unlikely that many homeowners in golf course neighborhoods actually rely on conservation easements because, generally speaking, whether a piece of land is encumbered by a conservation easement is not common knowledge.\textsuperscript{188} That is, if homeowners are unaware that the neighboring golf course is

\textsuperscript{182} See McCormick, supra note 181; Shaffer, supra note 181.
\textsuperscript{184} Ness, supra note 183. Contra Lowy, supra note 181 (“What kind of wildlife uses golf courses? The two that come to mind are earthworms and geese.”).
\textsuperscript{185} Ness, supra note 183.
\textsuperscript{186} See id.; see also Shaffer, supra note 181.
\textsuperscript{187} Kenneth A. Stahl, Reliance in Land Use Law, 2013 BYU L. REV. 949, 958 (2013) (“[C]ourts seem to think it fundamentally unfair that a landowner should expend significant resources on an investment in the good faith belief that the status quo would remain unchanged, only to endure a complete wipeout of that investment when an unpredictable change occurs.”).
encumbered by a conservation easement, then they cannot reasonably rely on the conservation easement to inform their expectations. Consequently, they have no basis for assuming that the land surrounding their property is insusceptible to change. In the rare scenario where homeowners are aware of and reliant on a conservation easement, courts can weigh the homeowners’ interests in protecting their settled expectations against factors such as the number of golfers making use of the golf course, the quantity of housing that can be built on the land, the housing demand in the area, and the availability of other sites for housing development in the neighborhood. Land is scarce, and to some extent, the protection of settled expectations must give way to a solution to California’s housing crisis.189

CONCLUSION

Society’s understanding of science, its cultural values, and its needs evolve endlessly. The laws that form the boundaries of society must evolve as well. Our perception of golf courses as nature and California’s need for housing have evolved on parallel routes. It is now clear that golf courses are not a source of nature worth preserving, and California is in dire need of housing. Yet, we continue to protect golf courses under the guise of environmental conservation and maintain impediments to housing development. This anomaly can be resolved by transforming golf courses into housing. However, golf course conservation easements prevent that resolution because they perpetually prevent development on golf courses. Accordingly, another change is needed—a change to the legal landscape of conservation easements. By enabling the termination of conservation easements on golf courses, unproductive land can be freed to help alleviate California’s housing shortage.

---

189 See Stahl, supra note 187, at 958–59 (considering that the courts’ protection of reliance interests is limited in order to allow for adaption to changing circumstances).
A Case for Protecting Youth from the Harmful Mental Effects of Social Media

Kaidyn McClure*

INTRODUCTION ........................................................................................................ 326
I. META: A CASE STUDY.......................................................................................... 330
   A. The Algorithm as a Business Strategy ......................................................... 330
   B. The Impact of the Algorithm ....................................................................... 331
   C. Problems with Leaving Regulation to the Platform or to Teen Users .......... 333
II. THE PROBLEM OF THE EXISTING LEGAL LANDSCAPE TO DETER SOCIAL MEDIA GIANTS ............................................................. 335
   A. Current Interpretation of Section 230(c)(1) of the Communications Decency Act Shields Social Media Businesses from Liability .................................. 335
   B. Challenges Applying Existing Tort Law to Social Media Algorithms ............ 339
   C. Congress’ Proposed Bill .............................................................................. 341
III. A ROADMAP TO PREVENT FUTURE HARM .............................................. 342
   A. Incorporate New Understanding of Algorithms into Interpretation of Section 230(c)(1) .............................................................. 343
   B. Impose a Duty of Care in Light of a New Understanding of Algorithms ....... 347
   C. State Legislature or Courts? Set the Parameters for Social Media Businesses ................................................................. 353
   D. Enact Law that Encourages Businesses to Play an Active Role in a Healthier World ............................................................... 355
CONCLUSION ............................................................................................................ 358

* J.D. Candidate, Expected May 2023, Chapman University Dale E. Fowler School of Law. Bachelor of Marketing, University of San Diego, 2019. A very special thank you to the incredible production team and staff editors on Chapman Law Review for their dedication to the journal. I am grateful to Professor Noyes for encouraging me when I presented this topic and for providing invaluable feedback throughout the writing process. And, to my mom and dad. Your goodness inspires me. Thank you for our thought bursts—I couldn’t have done this without you!
INTRODUCTION

For years, businesses have executed strategies to engage viewers with their products or services. Since engagement strategies can be manipulative, marketers must consider whether and when certain marketing tactics are unethical. Today, social media companies may have the same basic objective to engage their audiences, but their engagement strategies utilize artificial intelligence. To keep users engaged on social media, these platforms deploy algorithms that manipulate what the user views based on the user’s predicted interests. But the algorithm doesn’t just dictate what a user sees. It amplifies the user-generated content, meaning that, while the underlying content may be created by a human, the user’s experience of the content, or of reality, is mediated by the algorithm.

This amplification is harmful because it enables the platform to show an unprecedented amount of personalized content to the viewer, ultimately promoting a message to the viewer that targets and preys on the viewer’s vulnerabilities and insecurities. This harm is evidenced by social media’s strong association with a rise in mental health challenges, primarily among teenagers.

---

1 See, e.g., Dr. Sydney Ceruto, The Psychological Concept That Can Make You a More Effective Marketer, FORBES: LEADERSHIP (Feb. 19, 2020, 8:45 AM), http://www.forbes.com/sites/fortescoscoachescouncil/2020/02/19/the-psychological-concept-that-can-make-you-a-more-effective-marketer/?sh=56f41c3c481a4 (describing how brands use classical conditioning to “train” customers to think about and turn to their brand).
2 See id.
5 See id. at 6–7.
6 See id. at 7, 11–12; see also Swathi Sadagopan, Feedback Loops and Echo Chambers: How Algorithms Amplify Viewpoints, THE CONVERSATION (Feb. 4, 2019, 4:18 PM), http://theconversation.com/feedback-loops-and-echo-chambers-how-algorithms-amplify-viewpoints-107935 [http://perma.cc/YEV8-XXCZ] (describing that algorithmic amplification is “when some online content becomes popular at the expense of other viewpoints” and experience shows that users viewing “a lighter version of a topic” are then recommended “more hardcore content”).
9 Recent research demonstrates that increasing social media use is an important factor affecting adolescents’ mental health, and it particularly adversely impacts girls. See Lennart Raudsepp & Kristjan Kais, Longitudinal Associations Between Problematic
existing scientific research shows the strong association between social media use and a decline in teen mental health.\textsuperscript{10} Teens are devoting so much time and effort to social media use that it limits other social activities, which researchers have coined “problematic social media use” or “PSMU.”\textsuperscript{11} In 2021, the U.S. Surgeon General squarely addressed the impact of harmful social media messages on teen mental health in a public advisory, stating that “too often, young people are bombarded with messages . . . that erode their sense of self-worth—telling them they are not good looking enough, popular enough, smart enough, or rich enough.”\textsuperscript{12} Evidence of the connection between mental harm and social media is further represented by lawsuits brought by parents against social media platforms, such as one against Instagram, alleging that the addictive algorithm caused their daughters’ poor self-esteem and depression, ultimately leading to suicide.\textsuperscript{13}

Not only do independent studies and public voices emphasize this strong association, but internal research performed by one of the social media platforms itself—Facebook (also the owner of Instagram)—exemplified that Facebook use caused mental harm to teens.\textsuperscript{14} In September 2021, Frances Haugen, a former

---

\textsuperscript{10} See id.

\textsuperscript{11} There is an increasing number of adolescents experiencing adverse effects due to PSMU. Id. Evidence suggests that an increase in adolescent girls’ PSMU is related to an increase in depressive symptoms. Id. at 3.

\textsuperscript{12} U.S. P UB. H EALTH S ERV., P REVENTIVE M ED. R EPS. 1, 1, 3 (2019), http://www.sciencedirect.com/science/article/pii/S2211335519300993 [http://perma.cc/3W2B-DHPX]. Female adolescents with social media profiles have significantly higher levels of depressed mood and lower self-esteem compared to young females that do not have a social media profile. See id. at 1.


Facebook product manager, released a host of internal reports demonstrating that Facebook’s amplification algorithm, including its engagement-based ranking on Instagram, negatively affects teen mental health and well-being. The algorithm’s engagement-based ranking enables Instagram to present specific content to the viewer based on personal user data collected by the platform and then amplify the user’s preferences. Haugen’s testimony illustrates that the algorithm is harmful; for example, it leads children from innocuous topics, like healthy recipes, to anorexia-promoting content.

So, is a typically reasonable business objective—to engage users—still reasonable when it is set in the context of social media and achieved by deploying artificial intelligence that lacks any sense of moral consequence? Is it reasonable when the underlying strategy causes harm to teen users in the form of depression, suicide, anxiety, and other emotional disorders, and the platforms are aware of these harms?

The existing legal landscape is ill-equipped to provide relief to teens suffering from mental harm caused by the algorithms and to hold Facebook and other social media companies accountable for such mental harm. The circuit courts’ current interpretation of section 230 of the Communications Decency Act broadly immunizes these providers, even if they deploy algorithms. This interpretation rejects any possibility that certain algorithmic functions may take providers out of the purview of immunity. Additionally, existing tort jurisprudence does not address the issue of mental harm caused by algorithmic capabilities, so courts would have to extend tort law to provide relief to teens suffering from mental harm. As for government regulation, the House of Representatives and Senate proposed

15 See Hagey et al., supra note 14. While some allege that Ms. Haugen had a political motive to release internal company documents, she denied any partisan motivations. See id. Additionally, the information reported is not contested. See id. On the contrary, the greater controversy was that Facebook’s research into Instagram’s effects on teen girls was hidden from the public and even some company advisory board members. See id.
16 See Focusing on Testimony from a Facebook Whistleblower: Hearings to Examine Protecting Kids Online Before the Subcomm. on Consumer Prot., Prod. Safety, & Data Sec. of the S. Comm. on Com., Sci. & Transp., 117th Cong. 8 (2021) [hereinafter Focusing on Testimony from a Facebook Whistleblower] (statement of Frances Haugen, former Product Manager of Facebook Inc.).
17 See id. at 13, 35.
18 Id. at 8.
20 See 47 U.S.C. § 230; see also discussion infra Part II.A.
21 See 47 U.S.C. § 230; see also discussion infra Part II.A.
22 See discussion infra Part II.B.
bills to address the mental health crisis from social media and provide relief to teens, but progress is slow.\textsuperscript{23}

Neither government nor society anticipated the serious, harmful effects that excessive use of social media would have on teen mental health today.\textsuperscript{24} Teens cannot protect themselves from depression, anxiety, addiction, and other negative side effects of Instagram’s engagement-based algorithm because they cannot control the content that they view; rather, the algorithm does.\textsuperscript{25} On the one hand, there is a need to protect teenage users against the negative consequences of Instagram, to deter social media giants from knowingly developing harmful algorithms, and to prevent further harm to teens. On the other hand, there is a competing interest to ensure that social media businesses are not unduly regulated or disadvantaged by overly broad mandates.

This Note proposes a roadmap for two non-mutually exclusive solutions to the problem of a deficient legal landscape for mental harm caused by certain social media algorithms. Part I leads the discussion with a focus on Facebook and Instagram, by exploring Facebook’s business model and the various externalities of Instagram’s algorithm. Part II describes the problem, arising out of courts’ broad interpretation of section 230(c)(1), existing tort law, and Congress’ proposed bill. Part III synthesizes a new reading of section 230(c)(1) and suggests extending tort law to provide relief in conjunction with the proposed interpretation of section 230(c)(1). Part III also proposes a legislative solution to hold Facebook and other companies like it accountable for writing algorithms that cause mental harm, noting the advantages and disadvantages of a legislative approach.

\textsuperscript{23} See S. 2917, 117th Cong. (2021) (no action has been taken since the bill was introduced to the in the Senate in 2021); H.R. 5449, 117th Cong. (2021) (no action has been taken since the bill was referred to the Subcommittee on the Constitution, Civil Rights, and Civil Liberties on November 1, 2022).

\textsuperscript{24} See Hearing on “Algorithms and Amplification: How Social Media Platforms’ Design Choices Shape our Discourse and our Minds” Before the Subcomm. on Priv., Tech. & the L. of the S. Comm. on the Judiciary, 117th Cong. (2021) [hereinafter Social Media Design Discourse Hearing], http://www.judiciary.senate.gov/imo/media/doc/Harris%20Testimony.pdf [http://perma.cc/5WRE-6CV3] (statement of Tristan Harris, President and Co-Founder of Center for Humane Technology) (“We are raising entire generations of young people who will have come up under these exaggerated . . . mental health problems. . . . If this continues, we will see . . . more children with ADHD, more suicides and depression—deficits that are cultivated and exploited by [social media] platforms.”) (alteration in original).

\textsuperscript{25} See Focusing on Testimony from a Facebook Whistleblower, supra note 16, at 28. To increase the control that people have over their News Feed, Facebook enables users to “reject the personalized ranking algorithm altogether and instead view their feed chronologically, meaning that their [feed] simply shows them the most recent posts from their eligible sources of content in reverse chronological order.” See Social Media Design Discourse Hearing, supra note 24. But see Levin, supra note 7, at 17 (“[T]he social media platform is in the best (perhaps the only) position to control what users see, so solutions premised on the free-market ideal of individuals choosing what content they view are unrealistic . . . ”).
I. META: A CASE STUDY

A. The Algorithm as a Business Strategy

The risk of mental harm to social media users is exacerbated by the way algorithms are evolving and being utilized in the platform. Facebook did not use an algorithm at its inception in 2004; the platform was merely a collection of disconnected profiles. Facebook played a passive role in the user experience, allowing users to independently search for friends or strangers without any active input from Facebook. Thus, a user was largely in control of their experience. In 2009, Facebook introduced an algorithm that “determined the order of stories for each user” to display the most “juicy” posts near the top of the page. This straightforward ranking system helped users stay engaged on the platform without taking control from the user. By 2016, Facebook was joined by other social media platforms like Snapchat (owned by Snap, Inc.) and was forced to compete for the attention of young users. To keep from losing young users’ attention, Facebook used the algorithm to implement a user retention strategy to help users form meaningful social interactions. The algorithm executed this strategy by showing users the posts with greater comments and replies. These posts tended to be more extreme in nature, leading to adverse effects that perhaps were not anticipated. Today, Instagram deploys amplification algorithms, including engagement-based ranking. These algorithms bombard users with content the user wants to see based on the personal data collected. The danger is the development of feedback cycles, where teens are using Instagram to self-soothe, but then are exposed to more content that preys on their fears and insecurities.

---

26 Facebook’s CEO changed the company’s name to Meta Platforms, Inc. See Hagey et al., supra note 15. For clarity and consistency, I will refer to the company as Meta and to the platforms as Facebook and Instagram respectively throughout this Note.
28 See id.
30 See id.
31 See id.
32 See id.
33 See id.
34 See id.; see also Focusing on Testimony from a Facebook Whistleblower, supra note 16, at 8 (“[T]o be able to share fun photos of your kids with old friends, you must also be inundated with anger-driven virality.”).
35 See Focusing on Testimony from a Facebook Whistleblower, supra note 16, at 28.
36 See id.
37 See id.
engagement-based ranking system is different than the straightforward ranking system because it takes control away from how the user experiences the platform.38 Rather than allowing the user to experience content on the platform under their own volition, the amplification algorithm dictates how the user experiences the content, creating a greater risk of harm by preying on the user’s vulnerabilities without the user even realizing.

While an engagement-based algorithm poses more risk to users, it helps brands advertise to a highly active audience that is more likely to be interested in the advertisement.39 Instagram overwhelmingly helps small businesses by affording them the opportunity to reach millions of viewers at a low cost, an opportunity that would not exist without the algorithm’s capabilities.40 In 2020, the platform supported about 2 million monthly advertisers and over 25 million business accounts.41 Since Instagram’s service is funded by advertisers, Instagram is encouraged to deploy the engagement-based algorithm because it attracts more advertisers, and thus aggressively generates more revenue.42 The result is that users engage with more businesses on the platform.43 But the opportunity for harm forms when targeted messaging comes not from these advertisers, but from Instagram itself by promoting a specific message to the user that is perhaps unhealthy or dangerous to keep the user engaged.

B. The Impact of the Algorithm

Two aspects of social media platforms like Facebook and Instagram give rise to the risk of user harm: (1) a business model based on advertising revenue, and (2) the need to compete for engagement with competitors, such as Snapchat, Twitter, and

---

38 See id. at 8 (noting that users are self-identifying that they do not have control over their usage and that their usage is materially harming their health); see also Social Media Design Discourse Hearing, supra note 24, at 3.(statement of Monika Bickert, Vice President for Content Policy, Facebook) (trying to give more control back to users through various solutions).


41 See id.


43 See Wyatt, supra note 40.
TikTok. First, a revenue model based on third-party providers will inherently motivate a business to consider those providers’ interests.\(^4^4\) Thus, even though Instagram’s stated mission is “[t]o bring you closer to the people and things you love,”\(^4^5\) the means employed by Instagram are actually motivated to help third-party advertisers—which may involve bringing users closer to content with implicit harmful messaging from Instagram.\(^4^6\) Second, competition in the social media space makes it more difficult to keep users engaged.\(^4^7\) A solution that addicts users to the platform—such as deployment of an amplification algorithm—is good for advertisers because it promises more traction over their content, and keeps Meta in the game as a competitor. However, it is the algorithm’s addictive effect that contributes to users’ mental harm.\(^4^8\)

In 2019 and 2020, Facebook’s in-house analysts became aware of the intense social pressure, addiction, body image issues, eating disorders, anxiety, depression, and suicidal thoughts resulting from teen girls’ Facebook addiction.\(^4^9\) For eighteen months in 2019-2020, Facebook conducted a “teen mental-health deep dive” which included focus groups, online surveys, and diary studies.\(^5^0\) The research concluded that problems of mental health were specific to Instagram, coining an issue of “social comparison,” defined as a person’s assessment of their own value in relation to the attractiveness, wealth, and success of others.\(^5^1\) The large cause of social comparison is the algorithm’s curation of photos and videos on the Explore Page.\(^5^2\) A presentation posted to Facebook’s internal message board indicated that 32% of teen girls feel worse about their bodies after using Instagram, and 40% of teen boys


\(^{4^6}\) See supra notes 39–42 and accompanying text.

\(^{4^7}\) See supra note 31 and accompanying text.

\(^{4^8}\) See Force v. Facebook, Inc., 934 F.3d 53, 86–87 (2d Cir. 2019) (noting that the algorithms deployed by social platforms such as Facebook and Twitter are designed to keep users using, and such manipulation of news feeds influences users’ moods).

\(^{4^9}\) See 167 CONG. REC. S6759 (daily ed. Sept. 29, 2021) (statement of Sen. Marsha Blackburn). But see Hagey et al., supra note 14 (noting that Facebook invests billions of dollars to protect the safety of its users).

\(^{5^0}\) See 167 CONG. REC. S6759 (daily ed. Sept. 29, 2021) (statement of Sen. Marsha Blackburn); Wells et al., supra note 8.

\(^{5^1}\) See Wells et al., supra note 8.

\(^{5^2}\) See id.
experience[d] negative social comparison.”

However, Instagram researchers found it challenging to convince other colleagues of the gravity of the findings, who instead pointed to studies from the Oxford Internet Institute showing little correlation between social media use and depression.

In 2021, a teenager shared her story with the Wall Street Journal, explaining her belief that Instagram caused her eating disorder. She started using the platform when she was thirteen-years-old and was repeatedly bombarded by images of “perfect abs and women doing 100 burpees in 10 minutes.”

The harm that people experience from social media use can rise to clinical-level depression that requires treatment and can even extend to self-harm. In fact, a director for the eating-disorders program at Johns Hopkins Hospital expressed that she commonly hears from patients that their condition was caused by social media tips. For those vulnerable to negative emotional distress, Instagram escalates it.

C. Problems with Leaving Regulation to the Platform or to Teen Users

Hoping that either the social media platform will self-regulate or that teen users will regulate themselves is ineffective to protect teen mental health. For example, Facebook and Instagram cannot be trusted to prioritize mental health over user engagement goals because they’ve chosen to deploy an addictive algorithm despite awareness of the harmful effects. Facebook has acknowledged that the platform is a “sensory experience of communication that helps us connect to others, without having to
look away.”62 It may be difficult to understand the algorithm,63 but rather than take real steps to mitigate harms caused by the algorithm, platforms like Instagram merely warn users that services are provided “as is,” with no guarantee that they will work perfectly all the time.64 The algorithm’s unpredictability and lack of any moral sense, coupled with Facebook’s lack of motivation to protect teen health, does not lead towards improved mental health absent legal deterrence.

Additionally, despite Facebook’s attempt to help users improve their experience by allowing them to alter their account settings,65 teens are not making these changes because they are already addicted to the algorithm experience. Facebook’s own research showed that those struggling with the platform’s harmful psychological effects weren’t logging off, even if they wanted to, because they lacked the self-control.66 Some teens have shared that they often feel addicted and know that their mental health is deteriorating but are unable to stop themselves from using the application.67 Between 2009 and 2019, the number of high school students who experienced “persistent feelings of sadness or hopelessness” increased by more than ten percent.68 One could argue it’s unreasonable to require Facebook to protect users from the negative effects that result from the mere act of scrolling over content, even if that scrolling is excessive, and hold Facebook liable when it falls short. However, the addictive effect of the amplification algorithm may be as harmful to teen mental health as the addictive effect of nicotine is to teen physical health, and the public’s knowledge of tobacco’s harm necessitated federal legislation to reduce harm to teens.69 Moreover, studies show that

62 Cole F. Watson, Protecting Children in the Frontier of Surveillance Capitalism, 27 RICH. J. L. & TECH. 1, 23 (2021) (noting that the platform intends for users to “enter a mental state called the ‘machine zone’: a connection between user and device that invokes a ‘loss of self-awareness, automatic behavior, and a total rhythmic absorption carried along a wave of compulsion’”) (citing SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER 449–50 (2019)).
63 See Focusing on Testimony from a Facebook Whistleblower, supra note 16, at 22.
64 Terms of Use, supra note 45.
65 See supra note 25 and accompanying text.
66 See Wells et al., supra note 8.
67 See id.
teens implicitly trust social media companies. Thus, Facebook and other companies like it should be held accountable for harm caused to its users’ mental well-being because it is aware of the risk of serious harm and affirmatively contributes to it by deploying the amplification algorithm. We may never fully quantify the impact of social media on the communicative and behavioral development of teens. But leaving the choice to the user about whether to use the service is not much of a choice at all, since the service is free and more than a socially acceptable habit—it is a prerequisite of daily encounter.

II. THE PROBLEM OF THE EXISTING LEGAL LANDSCAPE TO DETER SOCIAL MEDIA GIANTS

The existing challenge is two-fold. First, the circuit courts’ current interpretation of section 230 of the Communications Decency Act (“CDA”) shields interactive computer service providers, like Facebook and Instagram, from liability for harm caused by its algorithms. Second, even if the courts reinterpret section 230 in a manner that does not put the function of algorithms within the scope of protection, the court must still extend the tort theory of negligent infliction of emotional distress to provide relief to teens that suffer mental distress, with or without any physical injury. While Congress has proposed a bill to address the issue of mental harm caused by social media companies, the language of the proposed bill imposes broad liability on these providers by providing relief for mental harms caused to teens by mere usage of the platform.

A. Current Interpretation of Section 230(c)(1) of the Communications Decency Act Shields Social Media Businesses from Liability

Congress enacted the “CDA” “to protect children from sexually explicit Internet content.” But since the public policy of the United States is to prevent “content regulation by the Federal Government of what is on the Internet,” section 230 was added as an amendment to the CDA “to maintain the robust nature of

---

70 See, e.g., Watson, supra note 62, at 24 (noting that teenagers presume that technological companies act in the user’s best interest).
71 See Terms of Use, supra note 45; see also Danielle Keats Citron, How to Fix Section 230, B.U. L. REV. (forthcoming 2022) (manuscript at 9) (on file with author) (describing the Internet’s “totalizing impact,” inextricable from daily life).
72 See discussion infra Part II.A.
75 Force, 934 F.3d at 78–79.
Internet communication and, accordingly, to keep government interference in the medium to a minimum.”76 The hope was that interactive computer service providers would “self-regulate” and “provide tools for parents to regulate.”77 Section 230(c)(1) immunizes interactive computer services against liability arising from content created by third-parties: “No provider . . . of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”78 An “interactive computer service” means any “information service, system, or access software provider that provides or enables computer access by multiple users to a computer server . . . .”79 A content provider is a “person or entity that “is responsible, in whole or in part, for the creation or development” of the content,”80 but a website provider “can be both a service provider and content provider.”81

Shortly after its enactment, in interpreting section 230, the Fourth Circuit stated that Congress’s objective was to immunize service providers from potential liability for messages republished by their services to prevent these service providers from severely restricting third-party messages.82 Since then, circuit courts have construed section 230(c)(1) broadly in favor of immunity.83

The Second Circuit created a three-part test to determine whether section 230(c) shields the defendant from civil liability.84 The defendant is immune from liability for state law claims if: (1) it is a “provider or user of an interactive computer service”; (2) the plaintiff’s claims treat the defendant as the publisher or speaker of content; and (3) that content is provided by a content provider other than the defendant interactive computer service.85 Social media companies like Facebook are considered interactive computer service providers (“providers”).86 The problem is that courts equate algorithmic functions as functions of a publisher of third-party content, satisfying the second and third elements to immunize the provider.87

77 See Force, 934 F.3d at 79.
79 Id. § 230(f)(2).
80 Id. § 230(f)(3).
81 Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157, 1162 (9th Cir. 2008).
83 See, e.g., Force v. Facebook, Inc., 934 F.3d 53, 64 (2d Cir. 2019).
84 See id.
85 Id.
86 See id.
87 See, e.g., id. at 67–68.
The case *Force v. Facebook* was the first to address the effect of Facebook’s algorithm on Facebook’s status as a publisher.⁸⁸ In *Force*, the Second Circuit determined that Facebook acted as a “publisher” within the meaning of section 230(c) when Facebook provided third-parties with a forum to communicate messages to interested parties.⁸⁹ The court did not believe that the algorithm changed the nature of Facebook’s role as a publisher because many of the algorithm’s functions like the “matchmaking” equated to editorial decisions that providers “have made since the early days of the Internet.”⁹⁰ The court implicitly classified Facebook’s algorithm as a “neutral tool[]” because it matches third-party content to users based on their preferences.⁹¹ To support this finding, the court cited to precedent which concluded that such neutral tools merely perform the job that is an inherent part of publishing: “organizing and displaying content exclusively provided by third parties.”⁹² The problem with such a conclusion is that, as Judge Katzmann pointed out in his dissent, the “majority . . . ‘cuts off all possibility for relief based on algorithms like Facebook’s, even if . . . future plaintiffs could prove a sufficient nexus between those algorithms and their injuries.’”⁹³ Certain algorithms, like Instagram’s amplification algorithm, are unlike ordinary editorial decisions; they do not merely determine where third-party content should appear on the site, who should see it, and in what form, as the Second Circuit suggests is the traditional result of editorial decision-making.⁹⁴ The court even pointed out that the algorithm’s capability goes beyond the capability of editorial decisions by presenting users with targeted content of more interest to them.⁹⁵

At the time section 230(c) was enacted, and later when *Force* was decided, the full extent of an algorithm’s capability was unknown. Control was an important underlying presumption motivating Congress’s decision to give broad protection to

---

⁸⁸ In *Force v. Facebook, Inc.*, users claimed that Facebook was civilly liable for aiding and abetting acts of international terrorism. *Id.* at 61. The plaintiffs argued that Facebook’s algorithm, exploiting user engagement to predict and show third-party content most likely to interest and engage the user, makes it so that Facebook is not a “publisher” within the meaning of section 230(c)(1) of the CDA. *Id.* at 65. The majority struck down their claim in finding that Facebook was immunized from liability under section 230. *See id.* at 68.

⁸⁹ *See id.* at 65.

⁹⁰ *See id.* at 66–67.

⁹¹ *See id.* at 66 (citing Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157, 1172 (9th Cir. 2008)).

⁹² *Id.*

⁹³ *See id.* at 77.

⁹⁴ *See id.*

provides under section 230(c), but modern users do not have a great degree of control over their experience with certain providers that deploy amplification-type algorithms.\textsuperscript{96} Congress also presumed that Internet services have “flourished, to the benefit of all Americans,”\textsuperscript{97} but the rise in mental health problems among teens contradicts Congress’s presumption that all Americans are benefitting. Thus, including harmful algorithms within the scope of section 230 immunity, as the courts have done, undermines the underlying presumptions of the defense.

Two years after Force, the Ninth Circuit took up the issue of the effect of algorithmic functions on Google’s status as a publisher in Gonzalez v. Google LLC.\textsuperscript{98} The court concluded that an algorithm that shows particular content to a user based on that user's inputs does not strip the provider of immunity as a publisher of third-party content.\textsuperscript{99} The court determined that by providing a neutral platform, not prompting the submission of certain content, and not determining the “types of content its algorithm[] would promote,” Google did nothing more than republish third-party content.\textsuperscript{100}

By viewing these recommendation capabilities as editorial functions, negligence claims based on the provider’s algorithm will continue to be dismissed under section 230.\textsuperscript{101} But as we better understand algorithms’ capabilities, a generalization that the algorithm does nothing more than help providers perform ordinary editorial decisions, as articulated by the majority in Force, does

\textsuperscript{96} See Force, 934 F.3d at 68; see also 47 U.S.C. § 230(a)(2) (providing immunity based on the presumption that the services “offer users a greater degree of control over the information they receive, as well as the potential for even greater control in the future . . . .”).

\textsuperscript{97} 47 U.S.C. § 230(a)(4).

\textsuperscript{98} In Gonzalez v. Google LLC, plaintiffs asserted that Google was not immune under the CDA for using computer algorithms to match and suggest content to users based on their viewing history. Specifically, they alleged that by recommending ISIS videos to users, Google assisted ISIS in spreading its message, going beyond its role as a publisher of third-party content. See Gonzalez v. Google LLC, 2 F.4th 871, 881 (9th Cir. 2021). The United States Supreme Court granted plaintiffs’ writ of certiorari and heard oral arguments on February 21, 2023. See Transcript of Oral Argument, Gonzalez v. Google LLC, 143 S.Ct. 762 (2023) (No. 21-1333), http://www.supremecourt.gov/oral_arguments/argument_transcripts/2022/21-1333_f2ag.pdf. During questioning, Justice Sotomayor stated “[T]here is a line at which affirmative action by an Internet provider should not get them protection under 230(c).” See id. at 97. Later Justice Gorsuch added “Is an algorithm always neutral? Don’t many [providers] seek to profit-maximize or promote their own products? Some might even prefer one point of view over another.” See id. at 101. Finally, Chief Justice Roberts commented to respondents that the third-party content appears “pursuant to the algorithms that [providers] have. And those algorithms must be targeted to something. And their targeting . . . is fairly called a recommendation, and that is [the providers’]. That’s not the provider of the underlying information.” See id. at 119.

\textsuperscript{99} See Gonzalez, 2 F.4th at 895.

\textsuperscript{100} See id.

not comport with reality. Providers act beyond the functions of publishers and play active roles in the user experience—they make and send curated messages to achieve effective targeted messaging for third-party advertisers. Courts should adopt an interpretation of section 230 that does not categorically treat all algorithmic functions as publishing functions. If Congress adopts a carve-out for harmful algorithms, plaintiffs can survive a section 230 immunity defense and seek recovery for mental harm caused by certain algorithms.

B. Challenges Applying Existing Tort Law to Social Media Algorithms

To provide a remedy for mental harm caused by certain social media algorithms, state courts must extend existing tort law, specifically under the theory of negligent infliction of emotional distress (“NIED”). Under existing law, the weight of a plaintiff’s burden varies from state to state depending on ‘the characterization of the elements that must be established to bring an NIED claim. In California, the plaintiff must establish the traditional tort elements of duty, breach of duty, causation, and damages. A duty’s existence depends on reasonably foreseeable risks of emotional injury and a weighing of policy considerations for and against liability. Additionally, the right to recover as a “direct victim” for emotional distress arises from the breach of a duty that is assumed by the defendant or imposed on the defendant as a matter of law, or that arises out of the defendant’s preexisting relationship with the plaintiff.

The issue as to whether a duty of care for algorithms exists or should exist remains open for courts to address. Today, section 230 theorizes a duty of care in the general social media context,

---

102 See, e.g., Focusing on Testimony from a Facebook Whistleblower, supra note 16, at 23–24 (calling attention to algorithmic biases and computer-driven content under amplification algorithms); see also Alina Glaubitz, How Should Liability be Attributed for Harms Caused by Biases in Artificial Intelligence? 13 (Apr. 29, 2021) (Senior Thesis, Yale Dep’t of Pol. Sci.) (noting that some algorithms can appear to be “facially neutral” when in reality they are discriminatory in application).

103 See discussion infra Part III.A.

104 See, e.g., Alina Glaubitz, How Should Liability be Attributed for Harms Caused by Biases in Artificial Intelligence? 13 (Apr. 29, 2021) (Senior Thesis, Yale Dep’t of Pol. Sci.) (noting that some algorithms can appear to be “facially neutral” when in reality they are discriminatory in application).

105 See discussion infra Part III.A.


107 See Molien, 616 P.2d at 816.
but limits the duty to moderation of illegal content. Additionally, courts have raised concerns about imposing a duty of care. The Ninth Circuit stated that “[n]o website could function if a duty of care was created when a website facilitates communication, in a content-neutral fashion, of its users’ content.” The plaintiff’s challenge, it seems, is to persuade the court to impose a duty of care on the interactive service provider to refrain from deploying algorithms that cause mental anguish. Since the original goal behind section 230 immunity was to protect minors from harmful material by incentivizing providers to block and screen such content, imposing a duty on providers to police their own actions, rather than the actions of third-parties, would continue to protect minors without chilling third-party speech. As it becomes more apparent that providers are, in fact, aware of the negative effects of their service’s algorithm on teens, an argument for the imposition of a duty of care for algorithms can create an avenue for redress while not imposing unreasonable burdens on providers. This Note addresses in Part III that the courts should impose a duty on social media companies to the extent they deploy amplification-type algorithms, given the foreseeable risk of mental harm caused to teens.

Another obstacle to bringing a successful NIED claim is establishing causation—that the algorithm caused the plaintiff’s mental harm. There is a great risk that social media litigation might mirror tobacco litigation. Tobacco litigation, under common law causes of action, was unsuccessful for over thirty years because the scientific evidence was insufficient to establish a causal link between tobacco and cancer. Although the scientific community recognizes the link between social media and mental harm, the evidence is still developing and social media businesses are downplaying the linkage.

108 See Glaubitz, supra note 102, at 29 (noting that social media platforms only have a duty to remove content that is prohibited by law).
110 Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093, 1101 (9th Cir. 2019).
111 See generally Artiglio v. Corning Inc., 957 P.2d 1313, 1318 (Cal. 1998) (noting that the existence of a duty to use due care toward an interest that enjoys legal protection against unintentional invasion is a threshold element of a cause of action for negligence).
113 See Turley, supra note 69, at 446.
114 See Wells et al., supra note 8. Adam Mosseri, Instagram head, reported that the app’s effects on teen health are likely “quite small” despite evidence showing that Instagram is damaging for many. Id.
The third challenge is the element of damages. Leaders in modern health recognize an array of mental health hardships that persist among teens, yet the common law contemplates only those of a sufficient severity that are chronic, or that are more traditionally understood as mental health disorders. Additionally, some states do not permit recovery for emotional distress alone without any accompanying physical injury. A handful of states, however, have turned the page, recognizing NIED as a means to recover for mental anguish without physical injury. In Rodrigues v. State, the Hawaii Supreme Court supported extension of the law by noting an important legal interest in protecting individual freedom from “the debilitating effect[s] mental distress may have on an individual’s capacity to carry on the functions of life.” In jurisdictions that recognize recovery under NIED for emotional distress alone, the court need only apply existing law in determining the damages element to a claim alleging mental harm caused by social media algorithms. Alternatively, if the state court has not modified the traditional rule requiring physical injury, the plaintiff must persuade the court to extend the law to impose a duty of care and allow the plaintiff to recover for mental harm unaccompanied by a physical injury.

C. Congress’ Proposed Bill

To address mental harm caused by social media, Congress proposed a bill in September 2021 to create a federal tort against social media companies. The purpose of the tort is limited to the

115 See Adolescent Mental Health, World Health Organization (Nov. 17, 2021), http://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health [https://perma.cc/MZ5Z-PDMG] (noting that depression, anxiety, and behavioral disorders are the leading causes of illness and disability among adolescents, and failure to address adolescent mental health conditions leads to impairment of physical and mental health in adulthood).
116 See, e.g., Jarrett v. Jones, 258 S.W.3d 442, 448 (Mo. 2008) (en banc) (requiring proof that emotional distress is medically diagnosable and of sufficient severity to be medically significant). But see McAllister v. Ha, 496 S.E.2d 577, 583 (N.C. 1998) (noting that emotional distress “means any emotional or mental disorder, such as . . . neurosis, psychosis, chronic depression . . . or any other type of severe and disabling emotional or mental condition which may be generally recognized and diagnosed by professionals . . . ”).
117 See, e.g., Kallstrom v. U.S., 43 P.3d 162, 165 (Alaska 2002) (requiring proof of physical injury to award damages for NIED since plaintiff’s case did not fall under Alaska’s two narrow exceptions); see also Anderson v. Scheffer, 752 P.2d 667, 669 (Kan. 1988) (emphasizing that a plaintiff cannot recover for emotional distress unless that distress results in actual physical injury, and headaches and insomnia are insufficient proof of physical injury).
118 See, e.g., Rodrigues v. State, 472 P.2d 509, 519–20 (Haw. 1970) (finding that traditional policy concerns limiting NIED to the establishment of physical injury are unpersuasive); see also Shuamber v. Henderson, 579 N.E.2d 452, 456 (Ind. 1991) (expanding NIED law to entitle a plaintiff to recover for emotional harm irrespective of whether the emotional harm arises out of or accompanies any physical injury); see also Molien v. Kaiser Found. Hosp., 616 P.2d 813, 820 (Cal. 1980) (en banc) (holding that the unqualified requirement of physical injury for NIED is no longer justifiable).
119 Rodrigues, 472 P.2d at 520.
deterrence of physical and mental harm caused to children less than sixteen years of age by social media companies. While imposing liability for harm caused to teenagers is beneficial to prevent harm to a vulnerable and targeted user group, the companies may actually be incentivized by the language of this regulation to bury their heads in the sand, avoiding liability by asserting lack of knowledge of the harmed user’s age. Moreover, the tort is not narrowly tailored to meet the root of the problem: the deployment of amplification-type algorithms. Instead, the tort imposes liability for harm caused merely by use. Since social media companies like Instagram and Facebook have the resources and knowhow to alter their platforms to provide more beneficial services to users, liability should be narrowly imposed for harm caused by detrimental capabilities of the algorithm, rather than broadly imposed for harm caused by mere usage.

III. A ROADMAP TO PREVENT FUTURE HARM

Two different routes may prevent social media companies from deploying harmful algorithms that cause mental harm to teens: a common law approach and a legislative approach. Under a common law approach, a plaintiff’s success on an NIED claim depends on two important variables: (1) whether the court is willing to adopt an interpretation of section 230(c)(1) that does not treat all algorithmic functions as the function of a publisher; and (2) whether the court is willing to extend tort law as needed to provide relief, including finding that social media companies owe a duty of care in algorithmic development. The alternative route to protect teen’ mental health is a legislative approach: Congress allowing the states to regulate under section 230. States could enact laws broad enough to target the harmful conduct—deployment of dangerous algorithms like amplification algorithms—yet impose a burden that is narrowly tailored to solve the problem, consistent with the need for a narrow tailoring of liability.

122 See id. (providing social media companies with an affirmative defense to the federal tort by assertion that the company took reasonable steps to ascertain the age of each user, or that the company did not know or had no reason to know of the user’s age).
123 As discussed in Part I, the issue of amplification algorithms stems from social media platforms’ third-party advertising revenue model. The business model is at the heart of the problem. If the platforms were less concerned with engaging users to third-party advertising, a shift away from amplification algorithms would be easier to make. Some scholars have proposed structural reforms as a means to reduce harms caused by the platforms. See Social Media Design Discourse Hearing, supra note 24 (statement of Tristan Harris, President and Co-Founder of Center for Humane Tech., proposing structural reforms for tech platforms’ incentives that would strengthen our capacity to solve problems like addiction and mental health problems).
124 See id.
125 See discussion supra Part II.A–B.
with Congress’s policy under section 230.\textsuperscript{127} States could even draft such laws in ways that would not implicate section 230 by not premising liability on whether the provider was acting as a publisher of third-party content. Since the providers would not be able to raise section 230 in response to the state law claim, a new interpretation of section 230 would not be necessary to ensure the success of a plaintiff’s claim under state law.

These two approaches are not mutually exclusive. However, the legislative approach is preferable because legislators can contemplate business interests along with societal interests to achieve the ultimate goal: preventing harm to teen mental health caused by social media platforms. Additionally, one state’s law can be adopted by various states over time to create uniformity. This will ultimately put pressure on social media companies to return to the drawing board to deploy safer algorithms that do not endanger teen mental health.

A. Incorporate New Understanding of Algorithms into Interpretation of Section 230(c)(1)

This Part III.A proposes an interpretation of section 230, as it applies to algorithms, inspired by the minority opinions in \textit{Force v. Facebook} and \textit{Gonzalez v. Google LLC}. Courts should adopt the following interpretation because a social media company becomes a form of provider-created content and is not exempt from liability under section 230 when it deploys an algorithm that enables it to use third-party content amplifying its own message to users to further its own goals.

In the dissent of \textit{Force}, Chief Judge Katzmann suggested that the section 230 does not protect Facebook from claims based on its suggestion algorithms because these claims do not inherently treat Facebook as the publisher of third-party content.\textsuperscript{128} To determine whether the claim inherently treats Facebook as the publisher of third-party content, the appropriate question is whether a plaintiff’s claim arises from a third-party’s information and whether that inquiry requires the court to view the provider as \textit{the} publisher of that third-party information.\textsuperscript{129} Even though a provider may publish third-party content, that provider’s liability is limited to the harmful function it performs; liability is not based on the provider’s identity.\textsuperscript{130} Chief Judge Katzmann seemed to recognize that the

\textsuperscript{127} See supra notes 74–77 and accompanying text.
\textsuperscript{128} Force v. Facebook, Inc., 934 F.3d 53, 82 (2d Cir. 2019) (Katzmann, C.J., concurring in part and dissenting in part).
\textsuperscript{129} See id. at 81.
\textsuperscript{130} See id. (citing Fed. Trade Comm’n v. LeadClick Media, LLC, 838 F.3d 158, 174 (2d
actions of the interactive computer service provider fall on a continuum, where the provider may be the publisher of third-party content on one end, and the third-party may be the publisher of their own content on the other end (when the provider transforms into the speaker of its own message by way of certain algorithms). If the provider transforms into the speaker of its own message, the provider is not the publisher of that information but rather a promoter of its own message. This is because, in this case, the provider is only using the third-party content to promote its message through the process of amplification. While Chief Judge Katzmann focused on harms that Facebook’s algorithm causes by connecting users, the same idea—that an algorithm enables a provider to play an “affirmative role” in causing harm—is pointedly applicable to mental health harms that the algorithm causes. It is the basis for arguing why social media platforms perform non-editorial functions when they deploy these algorithms and are thus not within the scope of section 230.

Similarly, a concurring opinion by Judge Berzon in Gonzalez v. Google LLC suggests that some algorithms enable providers to perform functions that are not within the scope of traditional publication. Judge Berzon of the Ninth Circuit explained why targeted recommendations and affirmative promotion of interactions among independent users are outside the scope of the traditional publication, and thus are not protected by section 230. Under her view, there is a difference between distributing content to anyone who engages with it and connecting users to specific content, treating the latter as more analogous to a direct marketer than to a publisher. Going a step further, Judge Gould, in his dissent, correctly points out that providers like Google and Facebook can act affirmatively through algorithms to repeatedly direct content to susceptible users, and when plaintiffs’ alleged harm is caused by such action, those allegations do not treat the provider as a publisher of the third-party content.

Cir. 2016) (noting that the CDA only bars lawsuits seeking to hold providers liable for exercising traditional editorial functions, such as deciding whether to publish, withdraw, or alter content)).

See id. at 76–77 (explaining, through a hypothetical, that it “strains the English language” to say that when the provider targets and recommends information to users, it is acting as the publisher of that information).

See id. at 77.

See Gonzalez v. Google LLC, 2 F.4th 871, 914, 920 (9th Cir. 2021) (Berzon, J., concurring).

See id. at 914.

Id. (“Traditional publication has never included selecting the news, opinion pieces, or classified ads to send each individual reader based on guesses as to their preferences and interests . . . .”) (alteration in original).

See id. at 921 (Gould, J., concurring in part and dissenting in part).
Synthesizing the foregoing opinions, an interactive computer service provider becomes a form of provider-created content and is thus not immune under section 230 when (1) the algorithm enables the provider to select third-party content to affirmatively promote its own message, to (2) targeted or susceptible users, and (3) the provider’s suggestions immerse the user in a universe of ideas that gives rise to the probability of harm.

Under the first factor, the question is whether the algorithm merely facilitates communication and content of others or enables the provider to actively communicate with users. Purely neutral search functions exemplify the former, and amplification algorithms, such as recommendation and social connectivity algorithms, exemplify the latter.137 Even though Facebook’s algorithm relies on and displays third-party user content,138 the anxiety and depression that may result from ordinary use of the platform is caused by the specific algorithm—the engagement-based ranking system—that synthesizes the user data to send a targeted message to the user.139 A claim containing this allegation does not inherently fault Facebook’s activity as the publisher of specific third-party content, but rather as the promoter of Facebook’s own message.140 The recent cases brought against providers involved third-party content that was itself harmful or offensive.141 Yet, for users suffering from the engagement-based ranking system, it may be the case where each piece of content, on its own and viewed independently, is not itself harmful or offensive.142 It is in these cases where it is more apparent that the provider plays an active role as a promoter of its own message, rather than as a passive arranger of content. For example, one photo of “how to lose weight” is reasonably not harmful, but impounding a user with similar media several times per day for endless days intensifies and magnifies a message, one that cannot be ignored or assuaged by the user, impacting the user’s overall

137 See id. at 917 (Berzon, J., concurring).
138 See id. at 914, 917.
139 See Focusing on Testimony from a Facebook Whistleblower, supra note 16, at 28.
140 See, e.g., Force v. Facebook, Inc., 934 F.3d 53, 83 (2d Cir. 2019) (Katzmann, J., concurring in part and dissenting in part).
141 See, e.g., Cohen v. Facebook, Inc., 252 F. Supp. 3d 140, 146 (E.D.N.Y. 2017) (involving harmful content from third-party terrorist organization); Force, 934 F.3d at 59 (involving harmful content from third-party terrorist organization); Gonzalez, 2 F.4th at 881 (involving harmful ISIS messaging and videos); Herrick v. Grindr, LLC, 306 F. Supp. 3d 579, 586–87, 589 (S.D.N.Y. 2018) (involving harmful third-party content: impersonating profiles).
142 See Allison Zakon, Optimized for Addiction: Extending Product Liability Concepts to Defectively Designed Social Media Algorithms and Overcoming the Communications Decency Act, 2020 Wis. L. REV. 1107, 1144 (2020) (recognizing the idea that the content itself is not harmful but rather the way it is shown to the user).
mental health. This supports Judge Katzmann’s conclusion that Facebook plays two roles as a service provider: the publisher of third-party content and the promoter of its own message to target the user based on statistical analysis of user information. The latter is not protected under section 230.

Under the second factor, the question is whether the algorithm acts on user-generated data. For example, Google (through YouTube), Facebook, and Twitter promote content to users who are susceptible to the harmful consequences of repeatedly viewing a subject of media. Suggesting content to users without any determination of user interest does not pose the same harm because the user is not as vulnerable to the provider’s message. To not protect interactive computer service providers merely because they suggest content would be detrimental to the service models that rely on advertising revenue. However, where the algorithm displays curated content to a user it has determined is engaged with the content, this aspect contributes to the dominating effect of the provider over the user and thus sets the stage for harm to occur.

Under the third factor, the question is whether the cumulative effect of suggestive content dominates the user experience. Where the algorithm enables the provider to interject its own message through its suggestive content, the provider may envelop the user, “immersing her in an entire universe filled with people, ideas, and events she may never have discovered on her own.” Facebook’s purpose is to build tools to help people connect. However, the current algorithm metrics do not put Facebook in the category of a passive service provider, providing the user with neutral features to build and maintain relationships with other users. On the contrary, Facebook is more like a promoter, interjecting a targeted viewpoint through the display of content that immerses the viewer with ideas that are not of the user’s own volition. This function, executed by the algorithm, is beyond the traditional editorial functions that section 230 immunizes. The interjection may be as simple as “you may be interested in viewing this content or connecting with these people,” but it is a message that the user would not have received on a platform deploying a “neutral” algorithm. Similarly, YouTube’s algorithm recalibrates

---

143 See Gonzalez, 2 F.4th at 921 (Gould, J., concurring in part and dissenting in part) (describing how a seemingly neutral algorithm amplifies messages).
144 See Force, 934 F.3d at 83 (Katzmann, J., concurring in part and dissenting in part).
145 See, e.g., Gonzalez, 2 F.4th at 921 (Gould, J., concurring in part and dissenting in part).
146 See id. at 917 (Bezon, J., concurring).
147 Force, 934 F.3d at 83 (Katzmann, J., concurring in part and dissenting in part).
148 Id.
149 See id.
the user’s existing interests to steer them toward new interests, often by displaying more divisive and extreme content.\footnote{See id. at 87.} In both situations, the provider dominates the user by purposefully intercepting third-party content to convey a targeted message by the provider for the purpose of achieving any number of the provider’s goals, like keeping users engaged on the platform for longer periods of time.

Adopting an interpretation that carves out certain algorithms from section 230’s protection does not stunt the beneficial growth of the internet. Rather, such an adoption would help prevent the harmful effects of Internet use that were not understood at the time of its enactment. As Chief Judge Katzmann pointed out in his dissent in \textit{Force}, where claims rest not on the content of the information but on the rules of the algorithm, the congressional intent of section 230 does not compel the judiciary to provide immunity.\footnote{See id. at 77.} Moreover, the suggested carve-out is itself narrow, and thus would still advance section 230’s aim at giving providers breathing space to grow.\footnote{See Gonzalez v. Google LLC, 2 F.4th 871, 921 (9th Cir. 2021).} By broadly immunizing providers, they are not incentivized to make their algorithms safer, despite knowledge of the harmful impact on users.\footnote{See id. at 920 (noting that a genuine factual issue exists as to whether social media companies are aware of the risks to the public stemming from content-generating algorithms).} Taking providers out of the purview of section 230 for deploying algorithms that fall within the narrow confines of the proposed factors would reasonably deter service providers from utilizing such algorithms and incentivize modifications to promote beneficial growth of the Internet, rather than plague users with emotional distress. Lastly, the narrow door would allow legitimate state law claims to be reviewed.\footnote{See, e.g., Herrick v. Grindr, LLC, 306 F. Supp. 3d 579, 593 (S.D.N.Y. 2018) (barring NIED under the current reading of the CDA).}

B. Impose a Duty of Care in Light of a New Understanding of Algorithms

By adopting this Note’s proposal that some algorithmic capabilities treat social media companies as promoters of their own messages rather than as publishers of third-party content, remedial courses of action—such as NIED—should survive an immunity defense under section 230 if two issues are resolved in favor of the plaintiff. At this point, the first issue is whether providers owe a duty of care to users for deploying harmful algorithms.\footnote{See Artiglio v. Corning Inc., 957 P.2d 1313, 1318 (Cal. 1998).} If answered affirmatively, the second issue is
whether the tortious conduct is framed such that the alleged duty does not treat the interactive computer service provider as a publisher or speaker of third-party content.\(^\text{156}\) To evade the purview of section 230, this Part III.B will discuss how to frame the tortious conduct for an NIED claim by analogizing to two recent cases involving social media companies defending against negligent design claims.

The court should impose a duty of care on the defendant (interactive computer service provider) when (1) a person suffers severe mental harm from use of a social media platform, (2) the harm is caused by the platform’s algorithm, and (3) the platform knew or should have known of the foreseeable risk of harm. To determine whether a duty of care exists, state courts consider various factors. For example, the California Supreme Court considers the following:

The major [considerations] are the foreseeability of harm to the plaintiff, the degree of certainty that the plaintiff suffered injury, the closeness of the connection between the defendant’s conduct and the injury suffered, the moral blame attached to the defendant’s conduct, the policy of preventing future harm, the extent of the burden to the defendant and consequences to the community of imposing a duty to exercise care with resulting liability for breach, and the availability, cost, and prevalence of insurance for the risk involved.\(^\text{157}\)

Applying these factors, a duty should be imposed on providers, like Facebook and Instagram, that deploy engagement-based ranking algorithms because there is a high risk of emotional distress and harm caused by such algorithms.\(^\text{158}\) The extent of the burden on the defendant is appropriately limited if the duty is triggered only when the interactive computer service provider knows or has reason to know of the risk of harm from use of its platform. For example, with the revelation of Haugen’s insights, it is evident that Facebook has knowledge of the harm posed by its conduct, yet it has not proposed a solution to prevent the harm. As for the consequences to the community for the imposition of a duty of care on social media companies, they likely weigh more in favor of imposition. If liability causes social media companies to rework algorithms to improve the user experience, we can help improve mental health for a generation of people currently suffering.\(^\text{159}\) Also, liability would likely incentivize healthy technological innovation in the context of social media

---

\(^{156}\) See Barnes v. Yahoo!, Inc., 570 F.3d 1096, 1102 (9th Cir. 2009) (noting that what matters is whether the claim “inherently requires the court to treat the defendant as the ‘publisher or speaker’ of content provided by another”).


\(^{158}\) See discussion supra Part I.B.

\(^{159}\) See supra note 24 and accompanying text.
rather than hinder it, or worse, promote innovation that does not consider mental wellness at all. One of the difficulties establishing the duty is the closeness of the connection between the algorithm and the injury. The degree of closeness is exemplified by answering whether modification of the algorithm would prevent the emotional distress, since this is the obligation that would be imposed on Facebook. It may be difficult for a plaintiff to establish that the risk of harm could be prevented by modifying the algorithm when the claim is against a social media company whose internal research is not publicized, or where the company’s knowledge of the risk is not publicly apparent. But according to Haugen, Facebook’s internal reports show that modifying the amplification algorithm would alleviate the harms caused to users, and outside studies tend to show that the risk could be prevented.

A criticism to imposing a duty on social media companies is that these social media companies may be encouraged to be less vigilant or proactive in conducting internal studies. This is problematic because social media companies possess the data, resources, and workforce to conduct accurate research efficiently, so they are in the best position to assess the quality of their service and its impact on users. The state legislature is thus likely the more appropriate forum to simultaneously (1) encourage social media businesses to study the use of their platforms and develop their algorithms in pursuit of healthier

---

160 In the context of AI development for autonomous vehicles, the prospect of tort liability could hinder innovation because the market is still developing. See, e.g., Andrew D. Selbst, Negligence and AI’s Human Users, 100 B.U. L. Rev. 1315, 1326 (2020). This economic concern is not as great for the social media industry because it is not as new of a market. See, e.g., Brian Dean, Instagram Demographic Statistics: How Many People Use Instagram in 2022?, BACKLINKO, http://backlinko.com/instagram-users [http://perma.cc/F8HB-PJWD] (last updated Jan. 5, 2022) (noting that about 500 million users around the world access Instagram daily).

161 On the one hand, a “tight causal nexus” between conduct and its consequences is fundamental to a fair assignment of liability; however, on the other hand, an economic theorist may argue “that the goals of tort law lie in optimal deterrence or efficient risk allocation.” See, e.g., Selbst, supra note 160, at 1321.


163 See Focusing on Testimony from a Facebook Whistleblower, supra note. 16, at 6.

164 Instagram uses the information it gathers to study its service and “collaborate with others on research to make [it] better and contribute to the well-being of [the] community.” Terms of Use, supra note 45.

165 Notably, Facebook does not make its research public, even for academics and lawmakers who have asked for it. Wells et al., supra note 8.
user experiences, and (2) hold these businesses accountable for mental harm caused to users. 166

Under Colorado state law, the court considers a different set of relevant factors and reserves consideration of any other relevant factors based on competing individual and societal interests implicated by the facts of the case. 167 In English v. Griffith, parents asserted an NIED claim against a woman for engaging in an argument with their son, allegedly causing their son such severe emotional distress to the point of causing him to take his life. 168 The Colorado Court of Appeals was asked to impose a duty on an individual not to cause another, who was known to be susceptible to emotional distress, to take his life. 169 The court did not find that the defendant owed a duty because the defendant could not “reasonably be expected to anticipate the mental health consequences that may flow from otherwise ordinary conduct such as the argument that allegedly occurred” in the case. 170 Under this line of reasoning, one might similarly argue that providers like Facebook and Instagram cannot reasonably be expected to foresee the mental health consequences that may flow from otherwise ordinary conduct—the use of social media—and therefore, a duty should not be imposed. However, unlike in Griffith, where the likelihood of injury resulting from the ordinary conduct was “extremely low,” 171 the likelihood of mental harm among teens caused by usage of social media tied to the amplification algorithm is high. Moreover, Griffith involved a defendant who was an individual, not a business entity. 172 Society may be more hesitant to burden individuals with legal duties to guard against mental harm. Conversely, society may have a greater interest in imposing a legal duty on a multibillion-dollar entity 173 that holds tremendous power over users, wields user trust, and knowingly

166 See, e.g., Levin, supra note 7, at 16–17 (noting that government interference is justified where platforms can use the algorithm to set the agenda in harmful ways without government parameters); see also discussion infra Part IV.C (describing additional advantages to a state legislative approach).

167 See English v. Griffith, 99 P.3d 90, 94 (Colo. App. 2004) (considering, for purposes of imposition of a legal duty: “(1) the risk involved; (2) the foreseeability of harm to others and likelihood of injury as weighed against the social utility of the actor’s conduct; (3) the magnitude of the burden of guarding against the injury or harm; and (4) the consequences of placing the burden on the actor”).

168 See id. at 92.

169 See id. at 94.

170 Id.

171 Id.

172 Id.

deports an algorithm that exploits users’ personal vulnerabilities to control their experience of the platform.

To defeat a section 230 defense to an NIED claim, the plaintiff must ensure that its allegations do not treat the provider as a publisher of third-party content but rather as a promoter of its own message.\(^\text{174}\) The following two recent cases exemplify the differences between the former and the latter. In *Doe v. Twitter*, two thirteen-year-olds were manipulated into providing pornographic videos to a third-party sex trafficker, and the videos were posted on Twitter a few years later.\(^\text{175}\) They asserted a state law claim based on negligent design, seeking to hold Twitter liable for enabling users to disseminate information quickly to large numbers of people, as well as for failing to deploy measures that prevent suspended users from opening new accounts and disseminating harmful content.\(^\text{176}\) The district court held that these allegations treated Twitter as a publisher protected by the CDA because “Twitter would have to alter the content posted by its users” to meet the obligation plaintiffs sought to impose.\(^\text{177}\) In reaching this conclusion, the court distinguished the allegations from those made in *Lemmon v. Snap*, where a negligent design claim was not barred by section 230.\(^\text{178}\)

In *Lemmon v. Snap*, the plaintiffs “were parents of two boys who were killed in a high-speed car accident.”\(^\text{179}\) They brought the action against Snap, Inc., the owner of Snapchat.\(^\text{180}\) The parents alleged that Snapchat’s “speed filter incentivized young drivers to drive at high speeds” and that Snapchat “was aware of the danger” of the filter from news articles and other accidents linked to Snapchat users’ high-speed snaps.\(^\text{181}\) In this case, the negligent design was not barred by section 230(c)(1) because the claim sought to hold Snapchat liable for its conduct as a manufacturer rather than as a publisher of third-party content.\(^\text{182}\) The primary reason for this conclusion was the fact that Snapchat could have “take[n] reasonable measures to design

\(^{174}\) Cf. *Lemmon v. Snap*, Inc., 995 F.3d 1085, 1092 (9th Cir. 2021) (finding that a negligent design lawsuit treats the social media company as a products manufacturer, and the duty underlying such claims differs from the duties of publishers as defined in the CDA); see also *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096, 1102 (9th Cir. 2009) (finding that a determination of whether a provider is a publisher protected by the CDA is based on “whether the duty that the plaintiff alleges the defendant violated derives from the defendant’s status or conduct as a ‘publisher or speaker’” (quoting 47 U.S.C. § 230(c)(1)).


\(^{176}\) Id. at 930.

\(^{177}\) Id.

\(^{178}\) Id.

\(^{179}\) Id. at 929 (describing the facts of *Lemmon v. Snap*).

\(^{180}\) Id.

\(^{181}\) Id.

\(^{182}\) See id. at 929–30.
a product more useful than it was foreseeably dangerous . . . without altering the content that Snapchat’s users generate.”

Since the speed filter was affirmatively created by Snapchat, the flaw was dependent on Snapchat’s actions, rather than any posting of third-party content.

An NIED claim against social media platforms in which plaintiffs allege that the algorithm, like engagement-based ranking, causes mental harm is more like the claim made in Lemmon and should withstand a section 230 defense where courts adopt this Note’s proposal for an algorithm carve-out. For example, an NIED claim against Instagram would seek to hold Instagram liable for its promotional action: targeting third-party content at users to send a message from Instagram meant to keep the user engaged on the platform which, as a result, harms the user. This framing of Instagram’s conduct does not treat Instagram’s duty as that of a publisher of third-party content within the scope of section 230 immunity because the alleged duty does not rest on any affirmative obligation to remove, alter, monitor, or edit third-party content. Rather, it is a duty to use reasonable care to refrain from writing algorithms that enable Instagram to send messages to targeted users that foreseeably cause mental distress. Like in Lemmon, where the allegations treated Snap as liable for its conduct as a manufacturer, an NIED claim alleging that Instagram acted unreasonably by failing to deploy a safer algorithm, given foreseeable risks of harm, holds Instagram liable for its conduct as a business deploying a harmful algorithm, rather than for its conduct as a publisher.

\[\text{感恩于前人的智慧和努力，我们能够继续前行。} \]
C. State Legislature or Courts? Set the Parameters for Social Media Businesses

For social media platforms that follow an advertising-based revenue model, maximizing revenue will naturally put third-party advertiser interests at the forefront of algorithm development. Absent an economic motivation to otherwise prioritize users’ mental health, teenagers are at the mercy of the platforms. Thus, passing legislation that incentivizes social media companies to turn their attention back to the users may be the most effective approach to protect teenage mental health and well-being, especially in a world where, for many, the thought of dissolution of social media is unimaginable.

One advantage to a state legislative approach, as opposed to a judicial approach, is that deterrence of harmful social media practices is wrapped up in complex policy questions that are best left to each state. Although social media and tobacco are uniquely similar in their addictive qualities targeting teens, regulating social media is more convoluted than tobacco regulation because social media can be positive, and it is largely good for small businesses and other stakeholders—including the workforce, supply chain of businesses, and other advertisers.

The citizens of every state may feel differently about the extent of the burden that should be imposed on social media companies. For example, some states may wish to impose liability only for harms caused to vulnerable user groups, like teenagers, which is a limitation that cannot be imposed under an NIED cause of action. Rather than asking courts to extend tort law and create a zone of liability without considering the public voice, states can enact more optimal solutions that reflect competing interests. Although social media businesses would face fifty different remedies from state legislation, the first state law will serve as the blueprint for other states. Moreover, any patchwork of laws and judgments that may result would not likely contort the national market any more than state common law courses of action.

190 See supra note 44 and accompanying text.
191 See, e.g., Levin, supra note 7, at 34 (stating that government intervention “could be used to create ethical rules and norms that apply to all social media platforms, combined with the means to enforce them”).
192 See, e.g., Jacqueline Tabas, How Nonprofits Can Use Social Media to Increase Donations and Boost Visibility, FORBES (Mar. 6, 2021, 09:00 AM), http://www.forbes.com/sites/allbusiness/2021/03/06/how-nonprofits-can-use-social-media-to-increase-donations-and-boost-visibility/?sh=5800326a2bb7 [http://perma.cc/35HP-82G5] (noting how social media helped nonprofits achieve fundraising goals). Notably, the problem with the tobacco crisis was that states could enact laws that “eliminated the core defense needed by the tobacco industry to defend itself.” Turley, supra note 69, at 472.
193 See supra notes 40–41 and accompanying text.
194 See Turley, supra note 69, at 468.
Another advantage to a state legislative approach is that neither re-interpretation nor reform of section 230 is necessary to allow relief to teens for mental harm; thus, a broad interpretation of section 230 in conjunction with a narrow state law may cohesively work to achieve, deter, and prevent future mental harm. Congress gave states implicit permission in section 230(e)(3) to enact law pertaining to interactive computer service providers, stating that “[n]othing in this section shall be construed to prevent any State from enforcing any State law that is consistent with this section.” As long as the enforcement action doesn’t conflict with or undermine section 230, states may address challenges of interactive computer service providers under their general police power. A broad reading of section 230 affords social media companies protection while states are granted latitude to protect users. This is important because, as Professor Jonathan Turley has noted, states have an “interest in private litigation” and, if capable, can “construct procedures that can act like legal speedtraps to capture wealth.” Since the states cannot enact law that is inconsistent with section 230, social media companies would still be protected from allegations of liability for conduct that is outside their control, like the posting of harmful content by a third-party. States will then be afforded the opportunity to enact law that holds these businesses accountable for conduct that is within their control, like algorithm development. A critique of this argument is that even legislation will struggle to effectively regulate platforms given the fast-paced development of technology and business operations. However, the nuances of technology and the harms it causes are more appropriately handled by the legislature—as opposed to courts—since the legislature can rewrite, repeal, and amend, and is not bound by precedent.

On the other hand, why not a federal legislative solution? Federal regulation would establish uniform liability, eliminating the burden on social media companies of sifting through state laws to ensure compliance. However, state legislatures are the appropriate forum to craft a creative solution for a national

---

197 See, e.g., Barbier v. Connolly, 113 U.S. 27, 31 (1884) (noting states’ power, termed as the “police power,” to “prescribe regulations to promote the health, peace, morals, education, and good order of the people, and to legislate so as to increase the industries of the state, develop its resources, and add to its wealth”).
198 Turley, supra note 69, at 471.
Additionally, if the states legislate to solve for mental harm by imposing liability for the deployment of harmful algorithms—even if the damages differ, the standard of causation differs, or the pool of plaintiffs eligible to take advantage of the law differs—the laws would nonetheless have the same effect on the businesses by pushing them to deploy less harmful algorithms. Lastly, providers like Facebook are already accustomed to navigating unique state laws, like data privacy laws, and they make changes to their business to comply with these laws because it is in their best interest. For example, California enacted a privacy law that gives Californians special privacy rights. The law applies to Internet providers that “operate in the state, collect personal data for commercial purposes[,] and meet other criteria” like generating revenue that exceeds a threshold. In response to the new legislation, many providers, like Microsoft, decided to “apply their changes to all users in the United States rather than give Californians special treatment.” Similarly, if providers were faced with a state law that imposed liability for deploying algorithms that harm teens who reside in the state, the providers could act in a manner that benefits all teen users.

D. Enact Law that Encourages Businesses to Play an Active Role in a Healthier World

If we accept the premise that some government intervention is necessary and desirable to ensure that all persons do, in fact, benefit from the use of the Internet, as Congress believed was already the case, then the question is how to intervene. Social media platforms can be designed to foster community safety, even with the help of algorithms. A law that is broad enough to meet today’s problem of mental harms arising from social media use and prevent the problem of advanced targeted messaging

---

199 See New State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting) (“It is one of the happy incidents of the federal system that a single courageous state may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country.”).


201 Id.

202 Id.


204 See Social Media Design Discourse Hearing, supra note 24 (statement of Joan Donovan, Ph.D. Research Director at Harvard Kennedy School’s Shorenstein Center on Media, Politics and Public Policy).

205 See, e.g., Sadagopan, supra note 95 (explaining how algorithms can draw inspiration from human intelligence to break harmful feedback loops).
tomorrow may be the best solution. This solution requires that we meet the root of the problem: algorithms.

A state law should be directed at the creators of algorithms to encompass interactive service providers, as well as businesses that do not satisfy Congress’s definition of an interactive computer service. The law should be articulated as follows: A creator of an algorithm shall be liable to any consumer who suffers bodily injury or harm to mental health when the consumer was less than twenty-years-old that is attributable, in whole or in part, to the individual’s use of technology that deploys a covered algorithm, where the creator of the algorithm knew or should have known of the risk of harm to the user. The term “creator of an algorithm” means an interactive computer service or other business that uses a covered algorithm to enhance a service or product provided to consumers. The term “consumer” means purchasers, users, patrons, and clients. The term “interactive computer service” has the meaning given to the term in section 230 of the Communications Decency Act (47 U.S.C. § 230). The term “covered algorithm” means reinforcement algorithms, amplification algorithms, and any other

---

206 See Watson, supra note 62, at 19 (noting how the Internet service provider, Google, has already introduced a new artificial intelligence that significantly improves clickthrough rate predictions); see also Glaubitz, supra note 102, at 6-7 (explaining the four generations of artificial intelligence and pointing out that engineers have only begun to develop the second generation).

207 See generally Adam Beam, Social Media Addiction Bill Fails in California Legislature, AP NEWS (Aug. 11, 2022), http://apnews.com/article/social-media-california-legislature-f5fd4e8ac90546c5066c3a685ab58b2b [http://perma.cc/AU8U-EC2G] (discussing the failure of a bill that would hold social media companies accountable for knowingly using features that cause addiction). Since software development is at the core of Instagram’s business, and the company has decided that the benefits of its algorithm outweigh the costs of harm, it is appropriate to hold it accountable for its intentional development and deployment of the algorithm. Additionally, limiting liability to the deployment of an algorithm not only narrows the scope, but it also accounts for future algorithm-caused harms known to businesses beyond the social media space.

208 47 U.S.C. § 230(f)(2) (defining "interactive computer service" as "any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server").

209 See, e.g., Terms of Use, supra note 45 (discussing how Instagram uses automated technologies to ensure the functionality and integrity of the service).


212 See Social Media Design Discourse Hearing, supra note 24 (statement of Joan Donovan, Ph.D. Research Director at Harvard Kennedy School’s Shorenstein Center on Media, Politics and Public Policy) (noting that reinforcement algorithms pattern the distribution of content based on user signals to reinforce user interests).

213 See Focusing on Testimony from a Facebook Whistleblower, supra note 16, at 28; see also Social Media Design Discourse Hearing, supra note 24 (statement of Monika Bickert, Vice President for Content Policy, Facebook) (noting that amplification algorithms use a personalized ranking process driven by user choices and actions to sort content).
algorithm that enables the creator to select third-party content to communicate its own message to targeted consumers. An individual who suffers bodily injury or harm to mental health that is attributable, in whole or in part, to the individual's use of technology that deploys a covered algorithm where the creator knew or should have known of the risk of harm to the user may bring a civil action against the creator in an appropriate State court of competent jurisdiction for compensatory damages or actual damages, punitive damages, and attorney's fees and costs. If the user shows that the user's mental harm is attributable to the algorithm, and the creator knew or should have known of the risk of harm, the burden shifts to the creator to show that it acted reasonably in the deployment of the algorithm.214

The language of such a law is advantageous for several reasons. First, it is broad enough to remedy mental health harms caused to teenagers using social media platforms, yet also include other harms.215 Second, by targeting algorithms, the law does not interfere with content moderation practices or regulation thereof—which currently stand amidst the crossfire of differing policy viewpoints.216 This is because incentivizing safe algorithmic development doesn't impact the flow of third-party content itself on the platforms. Third, the language solves for unknown future harms caused by harmful algorithms by deterrence and through ease of amendment. As we discover more about the types of algorithms that cause harm to users, the legislature could amend the definition of “covered algorithms” to remain relevant and effective. Fourth, the law is reasonably tailored in two ways: (1) it imposes liability only for harm caused to teenagers, a more vulnerable and targeted group; and (2) it also limits liability to knowledge or scienter of the provider, which is in accordance with the literal language of section 230.217 Fifth, the burden on businesses is also reasonable because it does not impose liability for mere usage of the technology, like

---

214 See generally Citron, supra note 71, at 16–20 (introducing a “reasonable steps approach” as one way to reform section 230(c)(1) to solve for harm caused to users by third-party content).

215 See, e.g., Glaubitz, supra note 102, at 29 (describing disparate impact caused by algorithms); see also Gonzalez v. Google LLC, 2 F.4th 871, 921 (9th Cir. 2021) (Gould, J., concurring) (describing political violence caused by algorithms).

216 See, e.g., Nina I. Brown & Jonathan Peters, Say This, Not That: Government Regulation and Control of Social Media, 68 SYRACUSE L. REV. 521, 541–42 (2018) (arguing that regulation of content moderation risks First Amendment violations). But see Citron, supra note 71, at 22 (arguing that leaving the Internet under current regulation actually “chills valuable speech”).

217 See Gonzalez, 2 F.4th at 920–21 (Gould, J., concurring) (arguing that the text of section 230 does not suggest immunizing providers from liability for serious harms knowingly caused by their conduct).
Congress’s bill proposes, but rather limits liability to the deployment of particular algorithms described in the law.

As for the affirmative defense of reasonableness, the creator must be able to point either to robust internal research that does not show an association between the creator’s algorithm and depression, anxiety, suicidal thoughts, or other emotional distress, or to steps it took to prevent such mental harm. Regarding the latter, the law should articulate examples of reasonable steps. For example, if the creator deploys an algorithm that manipulates a person toward a targeted message, the creator can show that it took reasonable steps by alerting the person—while the person was using the service—that he or she received a targeted message by the computer service from the choice of content displayed. The creator could also set up a system on the platform where users answer survey questions aimed at understanding user mental health, then regularly post findings to public bulletins on the platform. Liability is ultimately imposed if the algorithm’s creator fails to show that it acted reasonably. As Professor Danielle Citron points out in her argument for section 230 reform, a reasonableness approach is “valuable precisely because it is flexible.” This kind of burden-shifting law may be the best way to balance society’s interest in protecting teenage mental health and the market’s interest in connecting small business advertisers with an engaged audience—all while incentivizing businesses to innovate algorithms in a healthier direction.

CONCLUSION

Innovation and creativity drive the world of marketing and business. New strategies will be developed to help businesses reach more people faster and at the least expense. Artificial intelligence is one such proven strategy, yet its value to some businesses is at a great cost to teen mental health. As we continue to discover the potential of artificial intelligence for targeted marketing, questions of law and ethics must be at the forefront. Today, we face a teen mental health crisis, partly impacted by social media algorithms. Social media platforms are best suited to change the nature of their algorithms to reduce harm, but change is not on the horizon where business models are based on third-party advertising.

218 See discussion supra Part III.C.
219 See Citron, supra note 71, at 19–20 (noting that requiring businesses to show reasonableness pressures platforms to keep up with best practices and defend those practices in litigation, ultimately establishing industry standards “that have the force of law to back them up”).
The legal landscape is currently ill-equipped to help teens seek legal redress for mental harm caused by social media algorithms. But we must find a way to hold social media companies accountable for the harmful externalities of tech development and protect teens from ongoing mental harm. Courts should adopt an interpretation of section 230 of the Communications Decency Act that does not bar claims seeking relief for mental distress caused by harmful algorithms. The amplification algorithm—and others like it—executes user engagement strategies that treat the social media platform as a promoter, not as a publisher of third-party content. A duty of care should be imposed on social media companies for algorithm deployment because these providers are in the best position to deploy alternative, less harmful algorithms. Furthermore, severe harm to teen mental health outweighs any associated cost to advertisers. Beyond the court system, state legislatures can directly target the root of the problem—algorithms—with laws that balance competing stakeholder interests. A state legislative approach is probably favorable to a common law approach, since the legislature can craft unique laws that consider both society’s stance on the extent of regulation and the future of algorithm development in the context of targeted messaging. If robust protective measures guard the stairs of technological innovation, we can take big steps toward ensuring teen safety and improving the lives of many.
Physician False Claims Act Liability—The Circuit “Split” That Illustrates the Need for Health Courts

Nicole Rickerd*

INTRODUCTION ............................................................. 362
I. THE FALSE CLAIMS ACT CIRCUIT “SPLIT” ................. 364
   A. Background of the False Claims Act .................... 364
   B. Circuit “Split” in the Healthcare Context .......... 370
   C. The (Mis)perceived Difference Between the Third
      and Eleventh Circuits’ Rulings ....................... 375
II. RECOMMENDATIONS FOR REFORM ....................... 381
   A. Inaccurate Healthcare Rulings Have Led to
      Numerous Externalities Demonstrating the Need for
      Specialized Healthcare Courts ...................... 381
   B. Structure of Reform ...................................... 383
      1. Federal Healthcare Courts with Article III
         Review .............................................. 383
         a. Proposed Federal Healthcare Court Structure ... 383
         b. Specialty Healthcare Courts are Constitutional 386
   C. Benefits of Health Courts ................................. 393
      1. Expertise Would Ameliorate Accuracy and Efficiency
         Concerns ............................................ 393
      2. The FCA (and Other Federal Healthcare Litigation)
         Constitute a Large Enough Portion of Government
         Revenue to Financially Justify the Recommended
         Health Courts .................................... 394
      3. Expanding the Existing System to Create Health
         Courts Would Financially Benefit Consumers .... 396

* J.D. Candidate, expected May 2023, Chapman University Dale E. Fowler School of
  Law. B.A. Claremont McKenna College, 2019. I would like to thank Professors Stephanie
  Lascelles and Celestine McConville for their guidance and invaluable feedback during the
  writing process, as well as my friends and family, especially my mother, for their
  encouragement and support throughout law school.
CONCLUSION .............................................................................................................398

APPENDIX I: FCA CASES AND RECOVERIES ATTRIBUTABLE TO HEALTHCARE .................................................................................................399

INTRODUCTION

A man was diagnosed with lung cancer.1 His oncologists discovered the cancer metastasized, meaning it spread to his brain and bones.2 The patient received chemotherapy treatments.3 Chemotherapy has numerous side effects, one of which is febrile neutropenia, a fever resulting from a patient lacking in a type of white blood cells.4 Unfortunately, the patient developed febrile neutropenia approximately six months after his diagnosis.5 He was admitted to the hospital, where he was treated by the inpatient physicians.6 The inpatient physicians prescribed and administered broad spectrum antibiotics, which resolved the patient’s fever.7 Concerned the patient’s immunocompromised condition subjected him to greater risk of infection, the inpatient physicians discharged the patient to continue the course of antibiotics at home with daily follow-ups from a home health nurse.8 Two days later, the patient’s fever spiked, and he was readmitted to the hospital through the emergency department.9 At this point, the inpatient physicians reviewed the patient’s medical records and discovered the patient had spent more than half of the past six months in the hospital for treatment of complications from the chemotherapy.10 They consulted the patient’s oncologist, who insisted that aggressive chemotherapy remained the appropriate course of action for the patient.11 The inpatient team disagreed, and felt the chemotherapy not only diminished the patient’s quality of life, but was further shortening his already expected six-month prognosis.12

1 See David J. Casarett, When Doctors Disagree, 8 AM. J. ETHICS 571, 571 (2006).
2 See id.
3 See id.
5 See Casarett, supra note 1, at 571.
6 See id.
7 See id.
8 See id.
9 See id.
10 See id.
11 See id.
12 See id.
As the adage goes, “[m]edicine is a science of uncertainty and an art of probability.”\(^{13}\) The reality is, physicians often disagree with each other.\(^{14}\) Disagreement can arise in many instances – whether it be the result of multiple treating teams as in the above example,\(^{15}\) the hierarchical nature of the medical system,\(^{16}\) patients seeking second opinions,\(^{17}\) or medical decisions questioned by insurance companies\(^{18}\) or government reimbursement programs.\(^{19}\)

\(^{13}\) ROBERT BENNETT BEAN, SIR WILLIAM OSLER: APHORISMS FROM HIS BEDSIDE TEACHINGS AND WRITINGS 125 (William Bennett Bean ed., Henry Schuman, Inc. 1950).

\(^{14}\) One study found seventy-seven percent of second opinions obtained after an initial diagnosis resulted in changes in diagnoses, treatments, or treating physicians. See Miles Varn, Data Shows Second Opinions Can Change the Course of Your Healthcare, PINNACLECARE (Apr. 9, 2015), http://www.pinnaclecare.com/highlights/blog/data-shows-second-opinions-can-change-the-course-of-your-healthcare/ [http://perma.cc/8AA4-D8GT].

\(^{15}\) See, e.g., Casarett, supra note 1, at 572. While this Note’s introduction described an example of an initial aggressive approach recommendation being called into question by a later recommendation to pursue a conservative course of treatment, the inverse can also occur. See, e.g., Francis D. Moore, What To Do When Physicians Disagree: A Second Look at Second Opinion, 113 ARCHIVES SURGERY 1397, 1398 (1978). This journal describes an older man with severe hip pain whose family physician determined he was too old to undergo any kind of an operation. See id. The patient is later seen by a surgeon who is very familiar with total hip reconstruction and who tells him:

I think it would be wise for you to consider a total hip. There is a risk to it and a mortality somewhere around 1%, with infection a possibility in about 3%, in our own hands. Even though you are 82 years old, your brain, heart, and kidneys are all working well. You deserve some more painless physical activity in the years left to you. The risk seems small, but you have the operation if you want it. See id.

\(^{16}\) In the United States, medical students report to interns, who report to residents, who report to attendings. See Jennifer Whitlock, Resident vs. Attending Physician: What’s the Difference?, VERYWELLHEALTH (Aug. 11, 2022), http://www.verywellhealth.com/types-of-doctors-residents-interns-and-fellows-3157293 [http://perma.cc/DH4L-XZDK]. Sometimes medical practitioners disagree with their supervisors’ decisions. See, e.g., Alex Harding, I Was Confident in My Patient’s End-of-Life Care. Then My Senior Doctor Overruled Me, STAT NEWS (Apr. 18, 2017), http://www.statnews.com/2017/04/18/medical-resident-attending-physician-disagreement/ [http://perma.cc/6QMH-MWWU]. This article describes a scenario wherein a resident was working in the cardiac intensive care unit treating a critically ill man who appeared close to death with little hope of reversing his decline. See id. As the man’s condition continued to worsen, the resident determined that escalating treatment would be pointless and would conflict with the family’s stated wishes. See id. The resident presented the patient’s case during morning rounds to the attending physician, who, after examining the patient, delineated orders for aggressive treatment protocols. See id.

\(^{17}\) See, e.g., Varn, supra note 14. A recent Gallup poll reported that about thirty percent of Americans seek second opinions about issues related to health or proposed treatment. See id.

\(^{18}\) In Rollo v. Blue Cross/Blue Shield, Tishna Rollo needed an autologous bone marrow transplant with high dose chemotherapy to treat a Wilms’ tumor, which is a malignant kidney tumor. See Rollo v. Blue Cross/Blue Shield, No. 90-597, 1990 U.S. Dist. LEXIS 5376, at *1–3 (D.N.J. Mar. 22, 1990). Blue Cross denied coverage upon determining the procedure in question was considered “experimental,” and as such was specifically excluded from coverage. See id. at *8.

\(^{19}\) See, e.g., United States v. AseraCare, Inc., 938 F.3d 1278, 1281 (11th Cir. 2019).
The latter can implicate complex issues, such as when a disagreement in medical opinion can subject the treating physician to liability under fraud statutes.

In particular, the False Claims Act (“FCA”) often deals with questions of medical necessity that can result in disagreement between the treating physician and the plaintiff’s expert. Under the FCA, “any person who . . . knowingly presents, or causes to be presented, a false or fraudulent claim for payment or approval . . . is liable to the United States Government.” The question then becomes, when a plaintiff’s expert disagrees with the treating physician’s assessment, can the treating physician be held liable under the FCA for making a false or fraudulent claim? Part I of this Note provides a background of the FCA and explores an alleged circuit split on the issue, ultimately concluding that the disagreement is more of a misunderstanding than an actual split. Although the circuits treat physician liability under the FCA very similarly, one circuit’s mischaracterization of another circuit’s decision muddled the case law, promoting judicial misunderstanding of the FCA and raising concerns that a lack of expertise in healthcare issues amongst judges has left them unprepared to grapple with complex medical terminology. Part II argues that such judicial confusion suggests that the current structure of judicial review does not meet the needs of the healthcare community and should be tweaked to include initial reviews by specialized federal health courts that expand upon the existing Medicare system, with “expert” judges to properly adjudicate healthcare litigation, such as that arising under the FCA.

I. THE FALSE CLAIMS ACT CIRCUIT “SPLIT”

A. Background of the False Claims Act

Historically, the FCA was the first whistleblower law in the United States and remains one of the strongest existing whistleblower acts. Originally enacted in 1863 during Abraham

---

20 See, e.g., id. Indeed, the individuals who often initiate FCA qui tam actions are healthcare professionals who, through their employment, notice cause for concern in the treating physician’s medical necessity certification. See, e.g., Winter ex rel. United States v. Gardens Reg’l Hosp. & Med. Ctr., Inc., 953 F.3d 1108, 1112–16 (9th Cir. 2020).

21 37 U.S.C. § 3729(a)(1)(A). Of note, the government is not always the plaintiff in FCA cases. See 37 U.S.C. § 3730. Individuals may bring civil actions for FCA violations, and in such qui tam actions, the government has discretion to intervene or allow the individual to proceed as the plaintiff. See id. In either scenario, the government receives a percentage of the recovery. See id.

Lincoln’s presidency as a governmental tool to address issues of fraud during the Civil War, it is sometimes referred to as “Lincoln’s Law,” and has been amended multiple times since its passage. While the FCA was originally enacted to combat military-related fraud, it also targets fraudulent acts in the areas of healthcare fraud, defense contracting fraud, financial fraud, conflicts of interest, cyber fraud, procurement fraud, grant fraud, customs fraud, and disaster relief fraud.


Procedurally, the FCA can be a *qui tam* cause of action, meaning relators can file cases on behalf of the federal government. These *qui tam* plaintiffs, who are private citizens, sue on behalf of the government and assume a share of the recovery if victorious. The government has the option to intervene within a sixty-day period, during which time the *qui tam* complaint is sealed, and is required to complete an investigation into the validity of the complaint. Extensions to the sixty-day time period can be granted, and at the conclusion of the investigation, the government makes a determination on whether to intervene, with the *qui tam* relator assuming responsibility for the case if the government declines to proceed. Additionally, relators cannot proceed with their case if the government already possesses knowledge of the facts that form the basis of the case. Often individuals who initiate FCA *qui tam* actions in the medical context are healthcare professionals who, through their employment, discover a reason for concern in the treating physician’s medical necessity certification. Of note, the government itself can also initiate FCA lawsuits on its own without a relator.


29 See What is the False Claims Act?, supra note 22.


32 See id.

33 See United States ex rel. McKenzie v. Bellsouth Telecomm., 123 F.3d 935, 939 (6th Cir. 1997) (quoting United States ex rel. Taxpayers Against Fraud v. Gen. Elec., 41 F.3d 1032, 1035 (6th Cir. 1994)) (noting the original source exception means a relator “is unable to pursue the suit and collect a percentage of the recovery if the case is based upon information that has previously been made public or if the claim has already been filed by another”).

34 See, e.g., Winter ex rel. United States v. Gardens Reg’l Hosp. & Med. Ctr., Inc., 953 F.3d 1108, 1112–16 (9th Cir. 2020). As a measure of protection since initiation of these actions often ties to the relator’s employment, those who report FCA violations have recourse if terminated or adversely impacted as a result of coming forward. See Benjamin McCoy & Zac Arbitman, *Blowing the Whistle: A Primer on the False Claims Act*, THE TEMPLE 10-Q (2019), http://www2.law.temple.edu/10q/blowing-the-whistle-a-primer-on-the-false-claims-act/ [http://perma.cc/C7WL-K22R]. These individuals are entitled to reinstatement with seniority restored, twice their back pay with interest along with compensation for additional damages, See id.

Substantively, individuals prosecuted pursuant to the FCA for “knowingly present[ing], or caus[ing] to be presented, a false or fraudulent claim for payment or approval” can face civil penalties and treble damages. The FCA defines “knowing” and “knowingly” as actual knowledge, deliberate ignorance of truth or falsity, or reckless disregard of truth or falsity. Specific intent to defraud is not a requirement under the statute. The claim element of the Act can be satisfied by any request for money made to an agent of the United States. The claim need not be paid or approved, only submitted. The falsity requirement of the FCA is less straightforward, in no small part because the terms “false” and “fraudulent” remain undefined statutorily. In the healthcare context, this ambiguity gives rise to the question of whether and when courts can deem a physician’s opinion false.

While the FCA is applicable to all federally funded programs, in 2020, the federal government recovered over $1.8 billion in healthcare-related FCA cases, which represent over eighty percent of all FCA awards. Two major federally funded healthcare...
programs are Medicare and Medicaid. In 2020, there were over 62.8 million Medicare beneficiaries and 75.3 million Medicaid beneficiaries. Medicare coverage is limited to products and services deemed “reasonable and necessary” for diagnosis or treatment and within the scope of benefits. When a patient presents to a physician with Medicare or Medicaid coverage, the physician certifies the medical necessity to the government for reimbursement of services. A service is “reasonable and necessary” if it “meets, but does not exceed, the patient’s medical need,” and is “[f]urnished in accordance with accepted standards of medical practice for the diagnosis or treatment of the patient’s condition . . . in a setting appropriate to the patient’s medical needs and condition.” In layman’s terms, this means health care services “needed to diagnose or treat an illness, injury, condition, disease, or its symptoms and that meet accepted standards of medicine.”

Hospice care, a federally-funded Medicare benefit, is one program often the subject of FCA lawsuits. In fact, “51.6 percent of all Medicare decedents were enrolled in hospice at the time of death in 2019.” Similar to other Medicare certifications, when a

53 See, e.g., United States ex rel. Druding v. Druding, 952 F.3d 89, 91 (3d Cir. 2020).
Physician certifies a patient for hospice, the physician attests that the patient has six months or less to live.\textsuperscript{55} Beyond the general time-based guidelines, there are disease-specific guidelines that can be employed for hospice certification if the patient meets the specific criteria established for the disease in question.\textsuperscript{56} Patients diagnosed with diseases or conditions such as cancer, amyotrophic lateral sclerosis, dementia, heart disease, HIV, liver disease, pulmonary disease, renal disease, acute renal failure, chronic kidney disease, stroke, and coma can qualify for hospice certification upon meeting certain criteria.\textsuperscript{57} Additionally, hospice patients must have a Palliative Performance Scale\textsuperscript{58} below seventy percent and exhibit dependency on a minimum of two activities of daily living to qualify.\textsuperscript{59} Finally, qualification for hospice certification can be achieved by meeting the “Decline in Clinical Status Guidelines” or presenting with certain diagnoses such as brain, small cell, or pancreatic cancer.\textsuperscript{60} Despite this abundance of protocols and criteria, physicians’ original prognoses can still prove inaccurate.\textsuperscript{61}

\begin{itemize}
\item \textsuperscript{56} See By the Bay Health, Determining a Patient’s Prognosis of Six Months or Less for Hospice 1–2 (2021), http://bythebayhealth.org/wp-content/uploads/2021/05/determining-a-patients-prognosis-of-six-months-or-less-5-19-20.pdf [http://perma.cc/9G73-UFAN].
\item \textsuperscript{57} See id. at 2–12.
\item \textsuperscript{58} The Palliative Performance Scale is a tool used to measure functional performance of, and predict survival among palliative care patients, based on measurements of “ambulation, activity level and evidence of disease, self-care, oral intake, and level of consciousness.” See Dawon Baik et al., Using the Palliative Performance Scale to Estimate Survival for Patients at the End of Life: A Systematic Review of the Literature, 21 J. PALLIATIVE MED. 1651 (2018), http://www.ncbi.nlm.nih.gov/pmc/articles/PMC6211821/pdf/jpm.2018.0141.pdf.
\item \textsuperscript{59} See By the Bay Health, supra note 56, at 2. Activities of daily living are “essential and routine tasks that most young, healthy individuals can perform without assistance,” such as walking, feeding and dressing oneself, and bathroom care including personal hygiene, toileting, and continence. See Peter F. Edemekong et al., Activities of Daily Living, NAT’L LIBR. MED.: NAT’L CTR FOR BIOTECH. INFO. (Nov. 19, 2022), http://www.ncbi.nlm.nih.gov/books/NBK470404/ [http://perma.cc/4UJQ-YD6N].
\item \textsuperscript{60} See By the Bay Health, supra note 56, at 1–2, 12–14.
\item \textsuperscript{61} One study found 13.4% of hospice patients outlive their original prognosis. See Pamela S. Harris et al., Can Hospices Predict Which Patients Will Die Within Six Months?, J. PALLIATIVE MED. 894, 895 (2014), http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4118712/ [http://www.perma.cc/C67U-42DS]. It also found only 48.4% of stroke patients, 36.6% of dementia patients, and 89.1% of cancer patients died within the expected time frame. See id. Medicare figures noted hospice survival figures exceeding six months in 11.8% of patients in 2010 and 11.4% of patients in 2011. See id.
All such certifications are subject to the FCA. While guidelines exist to help physicians make these determinations, it often comes down to judgment calls. Unsurprisingly, physicians often disagree about these complex decisions. The question of how these disagreements should be treated under the FCA, namely whether and when a treating physician can be held liable for making a false statement if a plaintiff’s expert physician disagrees with the medical determination, has been of great interest to courts in recent years.

B. Circuit “Split” in the Healthcare Context

Lately, federal courts have explored the meaning of “false” within the healthcare context of the FCA. Some courts take a flexible approach to the potential for medical opinions constituting falsehoods under the FCA. The Third, Ninth, and Tenth Circuits have, in certain scenarios, found that differences in opinion qualify as “false” under the FCA. Other courts appear more stringent

66 Worth noting is the public policy concerns of over-subjecting physicians to liability when plaintiffs engage in “expert shopping” to find physicians willing to condemn the treating physician’s approach, even when most physicians would not. See, e.g., Alaw Gray, Expert Shopping: An Overview and Top Tips, HILL DICKINSON (June 28, 2018), http://www.hilldickinson.com/insights/articles/expert-shopping-overview-and-top-tips [http://www.perma.cc/PFC4-M9VQ] (discussing judicial discouragement of “expert shopping,” which notably only targets the practice of changing experts after retaining an expert, and does nothing to prevent vetting experts before retaining by looking at their history of favoring plaintiffs or defendants, or interviewing them to get a sense of how they might testify).
68 See United States ex rel. Druding v. Druding, 952 F.3d 89, 91 (3d Cir. 2020); Winter ex rel. United States v. Gardens Reg’l Hosp. & Med. Ctr., Inc., 953 F.3d 1108, 1112–13 (9th
about when an opinion can be false under the FCA. The Fourth and Seventh Circuits have required “objective falsehood” to establish falsity,69 and the Eleventh Circuit recently echoed that sentiment in the healthcare context when it held that a reasonable difference in medical opinion cannot constitute a false statement pursuant to the FCA.70

The Third Circuit appears to embrace a flexible approach in *United States ex rel. Druding v. Druding*, where it rejected the objective falsehood requirement for FCA falsity.71 In *Druding*, the defendant’s former employees (many of whom served on an interdisciplinary team of clinicians that conducted a bimonthly review of patients up for hospice recertification) initiated the FCA action, alleging that the defendant, a hospice-care provider, instructed its employees to inappropriately alter admitted patients’ Medicare certifications to reflect eligibility, when in truth those patients were ineligible for hospice care.72 Since hospice eligibility depends upon a patient having six months or less to live,73 the alleged falsehood here dealt with the accuracy of the patients’ prognoses.74 The pertinent evidence included two competing expert reports: one by the relators’ expert, and one by the defendant’s expert.75 The relators’ expert noted in his report “[d]etermining the prognosis of patients with a serious terminal illness referred to hospice is a difficult task that depends on the judgment and experience of clinicians and the consideration of

69 See United States *ex rel.* Wilson v. Kellogg Brown & Root, Inc., 525 F.3d 370, 377 (4th Cir. 2008) (finding defendants’ alleged representations about relatively vague maintenance provisions did not constitute objective falsehoods and accordingly could not establish a falsehood under the FCA); United States *ex rel.* Yannacopoulos v. Gen. Dynamics, 652 F.3d 818, 837 (7th Cir. 2011) (holding defendant manufacturers did not violate the FCA in a sale of fighter jets because there was insufficient evidence to prove the price was objectively false).

70 See United States *v.* AseraCare, Inc., 938 F.3d 1278, 1281 (11th Cir. 2019).

71 *United States ex rel. Druding*, 952 F.3d at 91.

72 See id. at 91–92.


74 See United States *ex rel. Druding*, 952 F.3d at 91.
survival evidence from the literature,” but went on to opine of the forty-seven patient records the expert reviewed, thirty-five percent of the defendant’s patients were inappropriately certified for hospice care. The defendant’s expert disagreed, testifying instead that a reasonable physician would have found each of the contested hospice certifications contained accurate attestations of those patients’ hospice eligibility. When the defendant moved for summary judgment, the district court granted the motion upon finding that the experts’ “diverging opinions did not create a genuine issue of material fact about the falsity of a physician’s determinations that the patient [met] hospice eligibility” without evidence of objective falsity. When the relators appealed the district court’s decision, the Third Circuit considered whether conflicting expert testimony could generate a genuine dispute regarding a Medicare claim’s falsity and found in the affirmative, even going so far as to explicitly reject the objective falsehood requirement for FCA falsity.

The Ninth Circuit also rejected the objective falsity standard. In Winter ex rel. United States v. Gardens Regional Hospital & Medical Center, Inc., the defendants’ former Director of Care Management accused them of falsely certifying to Medicare that patients’ inpatient hospitalizations proved medically necessary. In the course of her employment, relator noticed a trend of an unusually high number of patients from the defendant nursing home being admitted to the defendant hospital, and detailed sixty-five incidences of allegedly improper hospital admission that were certified to Medicare for reimbursement. Here, unlike Druding, the record did not yet contain any expert opinions, but merely the allegations in the complaint which included lack of support in the

---

77 See United States ex rel. Druding, 952 F.3d at 91.
78 See id.
79 Care Alts., Inc., 346 F. Supp. 3d at 688.
80 See United States ex rel. Druding, 952 F.3d at 91–92.
82 Relator determined these admissions did not meet defendant hospital’s admission criteria and were unsupported by the patients’ records. See id. at 111.

Admitting a patient to the hospital for inpatient—as opposed to outpatient—treatment requires a formal admission order from a doctor ‘who is knowledgeable about the patient’s hospital course, medical plan of care, and current condition.’ Inpatient admission ‘is generally appropriate for payment under Medicare Part A when the admitting physician expects the patient to require hospital care that crosses two midnights,’ but inpatient admission can also be appropriate under other circumstances if ‘supported by the medical record.’

Id. at 1113–14 (citations omitted).
medical records, when the district court granted defendants’ motions to dismiss for failure to state a claim, upon deeming determinations of medical necessity “subjective medical opinion[] that cannot be proven to be objectively false.”\textsuperscript{83} On appeal, the Ninth Circuit was unpersuaded by the district court’s rationale.\textsuperscript{84} The Ninth Circuit expressly rejected the “objective falsity” requirement, noting that Congress imposed no such constraint and that “[a] doctor, like anyone else, can express an opinion that he knows to be false, or that he makes in reckless disregard of its truth or falsity.”\textsuperscript{85}

The Tenth Circuit appears to be aligned with the Third and Ninth Circuits on the question of objective falsity, given all three circuits have found opinions to be false under the FCA. In United States ex rel. Polukoff v. St. Mark’s Hospital, the relator accused his coworker, a physician, of performing thousands of unnecessary heart surgeries he fraudulently certified to Medicare as medically necessary.\textsuperscript{86} The relator also sued the employing hospital for complicity in the physician’s scheme.\textsuperscript{87} The complaint alleged the physician “fully understands, but rejects, the standard of care” and describes the surgeries at issue as “preventative.”\textsuperscript{88} The defendants thereafter filed motions to dismiss.\textsuperscript{89} The district court granted the defendants’ motions, reasoning that a physician’s medical judgment cannot be false under the FCA.\textsuperscript{90} The Tenth Circuit reversed the district court’s dismissal.\textsuperscript{91} Unlike the district court, which found that the treating physician’s certification could not be false absent a regulation clarifying the conditions under which it will or will not reimburse a procedure, the appellate court agreed with the position articulated by the Government (as \textit{amici}), that “[a] Medicare claim is false if it is not reimbursable, and a Medicare claim is not reimbursable if the services provided were not medically necessary.”\textsuperscript{92} Accordingly, the Tenth Circuit concluded that while FCA liability must be predicated on an objectively verifiable fact, verification of that fact can rely on clinical judgments which are

\textsuperscript{83} Id. at 1116.
\textsuperscript{84} See id. at 1113.
\textsuperscript{85} Id. at 1113.
\textsuperscript{86} See United States ex rel. Polukoff v. St. Mark’s Hosp., 895 F.3d 730, 734 (10th Cir. 2018).
\textsuperscript{87} See id.
\textsuperscript{88} Id. at 737–38.
\textsuperscript{89} See id. at 739.
\textsuperscript{90} See id. at 740.
\textsuperscript{91} See id. at 746.
\textsuperscript{92} Id. at 739, 742.
vulnerable to proof of truth or falsity.93 Put succinctly, the court “did not create a bright-line rule that a medical judgment can never serve as the basis for an FCA claim.”94

The Third Circuit interpreted the Eleventh Circuit to embrace a different standard in United States v. AseraCare, Inc., which also addresses the potential falsity of hospice certifications.95 In AseraCare, Relators filed the qui tam FCA lawsuit against their former employers, operators of hospice facilities.96 The Government intervened, alleging defendants submitted documentation falsely certifying certain Medicare recipients as terminally ill, when the Government determined otherwise.97 Like Druding,98 the relevant evidence here was both parties’ expert testimony.99 The Government’s expert testified that out of 223 of defendants’ patients, he would only have concluded 100 of them were eligible for hospice.100 However, the Government’s expert did not stop there. He went on to clarify that his testimony solely reflected “his own clinical judgment based on his after-the-fact review of the supporting documentation.”101 He further conceded his inability to discuss whether a treating physician was wrong about their patient’s eligibility.102 He also declined to refute defendant’s expert’s testimony that the prognoses were accurate.103 The Government’s expert never testified that no reasonable doctor could have concluded at the time of certification the patients at issue were terminally ill.104 Moreover, as the proceedings progressed, the Government’s expert actually changed his opinion concerning some of the patients’ hospice eligibility.105 The district court sided with the defendants, granting their motion for summary judgment.106 On appeal, the Eleventh Circuit considered whether a physician’s clinical judgment that a patient is terminally ill can be deemed false “based merely on the existence of a reasonable difference of opinion between experts as to the accuracy of that prognosis.”107 The court

93 See id. at 742.
94 Id.
95 See United States v. AseraCare, Inc., 938 F.3d 1278, 1281 (11th Cir. 2019).
96 See id. at 1282, 1284.
97 See id. at 1284.
98 See United States ex rel. Druding v. Druding, 952 F.3d 89, 91 (3d Cir. 2020).
99 See AseraCare, 938 F.3d at 1285, 1287.
100 See id. at 1284–85.
101 Id. at 1287.
102 See id.
103 See id.
104 See id.
105 See id. at 1287–88.
106 Id. at 1281.
107 Id.
agreed with the district court, holding a battle of experts is insufficient to establish falsity.108

C. The (Mis)perceived Difference Between the Third and Eleventh Circuits’ Rulings

Since the AseraCare opinion, legal scholars have grappled with how to interpret “false” within the meaning of the FCA.109 Even the courts are disagreeing with each other’s rulings and engaging in statutory construction and congressional intent analyses to bolster their approaches.110 This debate has led to widespread perception of major differences between AseraCare, on the one hand, and the Third, Ninth, and Tenth Circuit rulings, on the other, when in fact no consequential distinctions exist – certainly nothing to constitute a split.111

In Druding, the Third Circuit specifically addressed the AseraCare ruling, finding the Eleventh Circuit also “determined that clinical judgments cannot be untrue.”112 The Druding court explicitly disagreed with AseraCare and claimed it “reach[ed] the opposite determination.”113 The Third Circuit interpreted the objective falsity standard as requiring a factual inaccuracy that can never be proven since opinions are subjective.114 The opinion then took a tangent, expressing concern that the AseraCare standard improperly conflated the statute’s falsity and scienter elements.115 The Third Circuit suggested that concerns about exposure of medical professionals to FCA liability whenever the Government procures an expert with a contrary opinion is better addressed solely through the scienter element.116 The Druding opinion turned to the Supreme Court’s analysis of false statements

---

108 See id. Of note, the district court’s grant of summary judgment in favor of defendant was vacated and remanded on a separate issue, namely the district court’s failure to consider the entirety of the evidence by constraining the Government to solely rely on the trial record. See id. at 1281, 1305.
111 See discussion infra Part I.C.
112 See United States ex rel. Druding, 952 F.3d at 100 (citing United States v. AseraCare, Inc., 938 F.3d 1278, 1297 (11th Cir. 2019)).
113 Id.
114 See id. at 95–97.
115 See id. at 96.
116 See id.
under securities laws, wherein it found an opinion can be considered false and establish liability under common law. The Third Circuit then held that since common law is the appropriate place to turn because Congress did not define “false,” opinions can be false under the FCA if the facts contained within the claim are untrue or the holder falsely certifies compliance with a statute or regulation that is a condition for Government reimbursement. Applying this theory, the Third Circuit concluded that “a difference of medical opinion is enough evidence to create a triable dispute of fact regarding FCA falsity,” and the Government need only prove the claim submitted as reimbursable was not in fact reimbursable to establish FCA falsehood.

The Third Circuit contends that the Eleventh Circuit’s AseraCare decision, which held a reasonable difference in medical opinion remains insufficient to subject a medical professional to FCA liability, is on the other end of the spectrum. The Eleventh Circuit found the underlying clinical judgment must reflect an objective falsehood to trigger FCA liability. The court further delineated this requirement:

Objective falsehood can be shown in a variety of ways. Where, for instance a certifying physician fails to review a patient’s medical records or otherwise familiarize himself with the patient’s condition before asserting that the patient is terminal, his ill-formed “clinical judgment” reflects an objective falsehood. The same is true where a plaintiff proves that a physician did not, in fact, subjectively believe that his patient was terminally ill at the time of certification. A claim may also reflect an objective falsehood when expert evidence proves that no reasonable physician could have concluded that a patient was terminally ill given the relevant medical records. In each of these examples, the clinical judgment on which the claim is based contains a flaw that can be demonstrated through verifiable facts.

The Eleventh Circuit contrasted objective falsehood with a reasonable difference of opinion, or in other words “[a] properly formed and sincerely held clinical judgment,” among physicians reviewing medical documentation after the fact, which is insufficient on its own to prove those judgments and associated

---

118 See id. at 95–97.
119 See id. at 96–97.
120 See id. at 97, 100.
121 See United States v. AseraCare, Inc., 938 F.3d 1278, 1281 (11th Cir. 2019).
122 See id. at 1296–97.
123 Id. at 1297 (emphasis added).
claims for reimbursement are false pursuant to the FCA.124 In arriving at the conclusion that an FCA claim fails as a matter of law if plaintiff neglects to prove anything beyond a mere reasonable difference of medical opinion, the Eleventh Circuit relied on the same Supreme Court precedent used by the Third Circuit to discredit the AseraCare ruling.125

Although the Third Circuit specifically singled out the Eleventh Circuit’s ruling in AseraCare, the rulings are not in conflict with one another. AseraCare subtly distinguished between reasonable and unreasonable.126 A careful application of that distinction to the different facts of the various cases elucidates a clear common denominator amongst the circuits – that reasonable differences in medical opinions can prove false.

The Third Circuit interpreted AseraCare to hold “that clinical judgments cannot be untrue.”127 Yet, this interpretation is not supported by the case itself.128 In fact, AseraCare specifically listed ways in which a medical provider’s judgment can be objectively false in the context of the FCA: where the medical provider (1) does not have a basis for the opinion due to failure to assess the patient’s medical records or condition, (2) does not actually believe the opinion asserted, or (3) comes to a conclusion no reasonable physician, nurse, etc., would have reached.129 To understand the AseraCare ruling—and its implicit agreement with Druding on the falsity standard—it is critical to closely parse the language and discern the difference between a reasonable and unreasonable medical opinion.130

In AseraCare, the Government’s expert disagreed with some of the treating physician’s certifications but did not find the treating physician’s determinations unreasonable.131 In contrast, the difference in opinion in Druding was not as clear cut. In Druding, relators’ expert did not make as many concessions and found certification inappropriate in a number of instances.132 Thus, Druding was a fitting case for the third type of objective falsity,
wherein relators were trying to prove an unreasonable difference in medical opinion. 133 Similarly, the physician in Polukoff, the aforementioned Tenth Circuit case, who certified unnecessary heart surgeries to Medicare for reimbursement, faced liability under the second theory of objective falsity because his concession that the surgeries were merely preventative showed that he never actually believed the surgeries were medically necessary. 134 Finally, although the Ninth Circuit also explicitly rejected the objective falsity requirement on the theory that physician's judgment is not insulated from liability, the facts of Winter fall under objective falsity, namely the first type wherein the treating physician lacked a basis for the opinion, because relator determined the admissions at issue did not meet defendant hospital's admission criteria and were not supported by the patients' records. 135

Indeed, the Ninth Circuit, which the Third Circuit perceived as aligned with it, aptly noted its ruling in Winter was not incongruous with AseraCare. 136 The Ninth Circuit correctly honed in on the reasonable and unreasonable distinction, explaining:

[T]he Eleventh Circuit was not asked whether a medical opinion could ever be false or fraudulent, but whether a reasonable disagreement between physicians, without more, was sufficient to prove falsity at summary judgment. (citation omitted) . . . [T]he court clearly did not consider all subjective statements—including medical opinions—to be incapable of falsity, and identified circumstances in which a medical opinion would be false. 137

In short, the Eleventh Circuit never asserted “clinical judgments cannot be untrue,” as the Third Circuit suggested and so vehemently disagreed with. 138

The Third Circuit led itself astray by accusing the Eleventh Circuit of conflating scienter and falsity in a case that did not implicate scienter at all. 139 Scienter is somewhat implicated in the second theory of objective falsity, wherein the certifying physician

133 See id.
136 See id. at 1118 (“The Eleventh Circuit’s recent decision in United States v. AseraCare, Inc. is not directly to the contrary.”) (citation omitted).
137 See id. at 1118–19 (citing United States v. AseraCare, Inc., 938 F.3d 1278, 1297–98 (11th Cir. 2019)).
138 See United States ex rel. Druding v. Druding, 952 F.3d at 100 (citing United States v. AseraCare, Inc., 938 F.3d 1278, 1297 (11th Cir. 2019)).
139 See id. at 95–96; see also United States v. AseraCare, Inc., 938 F.3d 1278, 1297 (11th Cir. 2019).
must not actually hold the asserted opinion, because that involves a physician knowing he or she is lying.\textsuperscript{140} Similarly, the first approach to objective falsity, namely lack of support for the opinion due to failure to examine or review medical records, would involve a knowing act because a physician would know if he or she neglected to familiarize him or herself with the patient.\textsuperscript{141} The interplay ends there. Approaching liability under the third objective falsity premise of reaching an unreasonable conclusion certainly does not implicate scienter.\textsuperscript{142} Nothing in the \textit{AseraCare} opinion implicitly required the physician to know his position was unreasonable; only that it must indeed be unreasonable.\textsuperscript{143} Therefore, the \textit{AseraCare} case, which falls under the third objective falsity premise, in no way implicated scienter.

The Third Circuit so engrossed itself with this irrelevant scienter analysis that it failed to notice the reasonable-unreasonable distinction in \textit{AseraCare}. Indeed, the Third Circuit contrasted \textit{AseraCare}'s conclusion that “[a] reasonable difference of opinion . . . is not sufficient on its own to suggest that those judgments . . . are false under the FCA” with its own conclusion that “a difference of medical opinion is enough evidence to create a triable dispute of fact regarding FCA falsity” and failed to realize the importance of the term “reasonable.”\textsuperscript{144} The Eleventh Circuit limited falsity to unreasonable differences of medical opinion.\textsuperscript{145} By omitting “reasonable” from its holding, the Third Circuit left open the possibility that both reasonable or unreasonable differences in medical opinion could be false under the FCA.\textsuperscript{146} Thus, on the core issue, the two circuits agree that unreasonable differences in medical opinion can be false.\textsuperscript{147} The only outlier is whether the Third Circuit also allows reasonable differences in medical opinion to constitute falsity under the FCA—an absurd premise once one considers the extraordinary liability physicians would face whenever exercising clinical judgment in any situation not purely black and white.\textsuperscript{148} In short, all circuit courts that have addressed physician liability under the FCA treat objective falsity very similarly.

\textsuperscript{140} See \textit{AseraCare}, 938 F.3d at 1297.
\textsuperscript{141} See \textit{id}.
\textsuperscript{142} See \textit{id}.
\textsuperscript{143} See \textit{id}.
\textsuperscript{144} \textit{United States ex rel. Druding}, 952 F.3d at 89, 100 (quoting \textit{AseraCare}, 938 F.3d at 1297).
\textsuperscript{145} See \textit{AseraCare}, 938 F.3d at 1297.
\textsuperscript{146} See \textit{United States ex rel. Druding}, 952 F.3d at 100.
\textsuperscript{147} See \textit{id}.
\textsuperscript{148} See \textit{id}. 
The circuit courts failed to recognize that they each reached the same basic conclusion—that opinions might be a basis for false claims. This perceived circuit split where none exists is not concerning in and of itself. Rather, it is a symptom of the disease, namely a widespread mishandling of health law by the courts deserving of attention. By overlooking the reasonableness requirement, the Druding ruling muddles case law by (1) engaging in a confusing analysis culminating in a tangent about scienter, and (2) leaving open the possibility of subjecting physicians to FCA liability for reasonable differences in medical judgment, which poses obvious public policy concerns.

The fact that the Third Circuit overlooked the reasonableness requirement at least raises the question of whether federal courts of general jurisdiction are prepared to handle the complexities of healthcare law. Lawyers and judges confront the reasonable person standard in many areas of law, from contracts, to torts, to criminal law. The Third Circuit missed this analysis in the FCA context because the reasonable

---

149 It is worth noting the Third Circuit’s misconstruction of the Eleventh Circuit’s reasoning is but one example of the extraordinary complexity involved in applying legal concepts in the context of healthcare. Courts have misinterpreted medicine in a variety of areas, not just during adjudication of FCA claims:

[M]isleading statements about medical realities are not uncommon when judges make medical decisions. I also claim that the result of such misleading statements by judges is costly. The credibility of the courts is undermined in the eyes of the medical profession, and the credibility of the medical profession is undermined in the eyes of the public. The result is greater public distrust of both law and medicine. A loss of faith in both professions is the result of the vicious circle of counterproductive moves set in motion by these flawed decisions.


150 See United States ex rel. Druding, 952 F.3d at 89, 95–96 (3d Cir. 2020).

151 See id. at 100.

152 See RESTATEMENT (SECOND) OF CONTRACTS § 43 cmt. d (AM. L. INST. 1981) (“The basic standard to which the offeree is held [in determining the legitimacy of an offeror’s indirect revocation] is that of a reasonable person acting in good faith.”).

153 See RESTATEMENT (SECOND) OF TORTS § 46, cmt. j (AM. L. INST. 1965) (“The law intervenes only where the distress inflicted is so severe that no reasonable [person] could be expected to endure it.”); see also RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 7(a) (AM. L. INST. 2010) (noting the tort for negligence imposes upon actors a duty of reasonable care).

154 See, e.g., People v. Hurtado, 63 Cal. 288, 292 (1883) (holding murder is reduced to manslaughter “when it is committed under the influence of passion caused by an insult or provocation sufficient to excite an irresistible passion in a reasonable person”).
doctor standard is not quite as conspicuous. For example, determining whether someone acted reasonably in failing to put up wet floor signs near wet, slippery stairs\textsuperscript{155} involves drawing on personal experience common to most individuals, whereas ascertaining whether a doctor formed a reasonable conclusion regarding a hospice certification, which involves consideration of numerous complex medical factors,\textsuperscript{156} is not so easily decided by someone without medical knowledge. Indeed, \textit{Druding} and \textit{AseraCare} dealt with these hospice factors.\textsuperscript{157} Moreover, before judges can even hope to weigh complex medical factors such as those involved in a hospice certification, they need to learn the corresponding medical terminology. Understanding medicine requires fluency in terminology unfamiliar to the average individual.\textsuperscript{158} Learning medical terminology is akin to learning a foreign language—there are whole dictionaries dedicated to the subject.\textsuperscript{159} When medical terminology becomes inextricably intertwined with legal concepts, such as the reasonable doctor analysis in the FCA context, the legal principals themselves also become muddled, resulting in erroneous opinions. This explains why the Third Circuit took a wrong tangent and accidentally overlooked the reasonable-unreasonable distinction entirely.

\section*{II. RECOMMENDATIONS FOR REFORM}

\subsection*{A. Inaccurate Healthcare Rulings Have Led to Numerous Externalities Demonstrating the Need for Specialized Healthcare Courts}

The lack of nuanced medical understanding in legal opinions, such as the above FCA rulings, has led to confusion in the legal and medical communities about when liability is imposed on medical practitioners,\textsuperscript{160} and has created a risk of imposing liability where none should exist.\textsuperscript{161} These outcomes

\footnotesize

\textsuperscript{155} See, e.g., Galef v. Univ. of Colo., 2022 COA 91, ¶ 4.
\textsuperscript{156} \textit{See Determining a Patient’s Prognosis of Six Months or Less for Hospice, supra note 56–60 and accompanying text; see also supra text accompanying notes 56–61.}
\textsuperscript{157} \textit{See United States ex rel. Druding v. Druding, 952 F.3d 89, 91 (3d Cir. 2020); see United States v. AseraCare, Inc., 938 F.3d 1278, 1281 (11th Cir. 2019).}
\textsuperscript{160} \textit{See discussion supra Part I.B regarding alleged FCA “circuit split.”}
\textsuperscript{161} \textit{See supra note 156 and accompanying text.}
have fostered discontent within the medical community. The 2017 reform objectives from the AMA include a goal to “reduce regulatory burdens that detract from patient care and increase costs,” an objective that the increased efficiency offered by health courts could further. Many of the proposals for health courts in the arena of medical malpractice are for the state level. However, the same arguments that can be made for state health courts, such as to avoid defensive medicine and promote efficient ruling to remedy court congestion, can also be made at the federal level, especially since medical practitioners are defending their professional choices both when facing a state lawsuit for medical malpractice or a federal lawsuit for violation of the FCA.

Moreover, since at least the 1960’s, issues surrounding overburdened federal courts have existed due to the burgeoning volume and complexity of cases channeled into the system. Fast forward nearly another thirty years, and Congress continues to examine the issue of clogged courts caused by “overwhelming caseloads, substantial litigation delays and spiraling costs.” The Third Circuit, at a minimum, aggravated this backlog by wasting resources in investing time into a belabored analysis of an


163 See, e.g., Peters, supra note 163, at 228 (discussing moving medical malpractice cases out of civil courts).


165 See, e.g., Peters, supra note 163, at 228 (discussing moving medical malpractice cases out of civil courts).


inconsequential scienter tangent and promoted further delay for future courts attempting to grapple with the ruling that misconstrued the basic underlying law in the process.  

B. Structure of Reform

1. Federal Healthcare Courts with Article III Review

   a. Proposed Federal Healthcare Court Structure

   Congress should designate Medicare administrative law judges and Appeals Council as generalized federal healthcare courts, expand their purview to address all civil federal health law disputes, including the FCA, and add judges as needed for caseload management. Medicare uses administrative law judges and a Medicare Appeals Council to make determinations regarding authorization or payment for healthcare, the amount health plans require enrollees to pay, and limits on quantity of items or services. Specifically, Medicare determinations are appealable as follows: (1) redetermination by a Medicare Administrative Contractor; (2) reconsideration by a Qualified Independent Contractor; (3) hearing before an administrative law judge; (4) review by the Medicare Appeals Council; and (5) judicial review in a United States District Court. The federal health courts or federal health administrative agency proposed in this Note should thus be an expansion of this program to encompass all Medicare and Medicaid lawsuits, including those related to the FCA and other fraud statutes. The healthcare cases contemplated by this Note would begin at the third stage in a hearing before an administrative law judge, then progress through the appellate structure. The benefits of this small subset of non-FCA Medicare disputes already being addressed in an administrative agency is three-fold. First, it decreases the cost of getting a new system up and running since some logistics are already in place. While the existing Medicare Appeals Council houses judges in eleven field

---

171 See United States ex rel. Druding v. Druding, 952 F.3d 89, 96 (3d Cir. 2020) (fixating on scienter).


offices, the existing pool of health expert judges will decrease the costs significantly. Even if Congress were to establish federal healthcare courts in every state, which is not necessarily required, each judge already in existence would save approximately $900,000. Moreover, these judges are already experienced in Medicare issues and accustomed to weighing evidence from medical experts, texts, and research. As such, these judges could identify difficult issues for people from a non-healthcare background to understand, which could then be the focus of a training program for any additional judges for the federal healthcare courts. Second, it bolsters the proof of the need for specialized courts and the presence of a sufficient number of cases to justify them. Namely, the creation of the Medicare appeals process indicates the traditional court system could not, on its own, handle adjudication of such cases. Third, Medicare and the FCA are both federal healthcare statutes, and the Medicare Appeals Council constitutionally presiding over Medicare appeals implies that piggybacking off that same system to augment the caseload with similar litigation would also be constitutional.


175 The Office of Medicare Hearings and Appeals also makes use of trained mediators to lessen the workload for ALJ teams. See HHS PRIMER: THE MEDICARE APPEALS PROCESS, supra note 173, at 88.

176 Although this Note justifies establishing health courts in each district, similar to the bankruptcy court system, it is worth noting this may go above and beyond what is necessary—should Congress create health courts as legislative courts, it might be able to do so by merely establishing a centralized federal health court system in Washington, D.C., similar to Tax Court. See MARK DESGROSSEILLIERS, PERSONAL JURISDICTION IN BANKRUPTCY CASES: YOU’VE GOT MAIL 8 (The Federal Lawyer, 2019) (“The Supreme Court has not, to date, directly decided the extent to which the Fifth Amendment might impose limits on a federal court’s exercise of personal jurisdiction over an out-of-state defendant in cases involving federal questions, including but not limited to bankruptcy-related matters.”).

177 See Madison Alder, Congress Weighs First District Court Expansion Since 1990 (1), BLOOMBERG L. (Aug. 9, 2021, 10:37 AM), http://news.bloomberglaw.com/us-law-week/congress-weighs-district-judge-bills-after-decades-of-inaction [http://perma.cc/Y578-SH9L] (“It costs roughly $900,000 to add a new judgeship. That accounts for salary, benefits, staff, equipment, and travel, but doesn’t include the cost of additional space or security.”). Using an existing system with judges already in place that can simply expand their caseload to accommodate FCA and other healthcare cases will mean adding fewer judgeships than creating a whole new system.


180 See 42 C.F.R. § 484.10 (2012).
Specialized review at both the trial and appellate levels is necessary because the struggle to understand the complexities of medicine affects both trial and appellate judges. While the FCA analysis above focused on the appellate courts’ confusion, the underlying district court ruling in Druding distorted caselaw, potentially contributing to the Third Circuit’s confusion. However, the Third Circuit’s misunderstanding cannot be entirely attributed to the district court’s distortion, especially since it rejected the district court’s interpretation of the caselaw and independently came to a different conclusion, opposite to that of the district court. Since medical misunderstanding pervades trial and appellate courts, a second layer of specialized review is necessary to ensure the medical-legal analysis is fully fleshed out and persuasive when a healthcare case reaches a non-expert review by a district court.

Administrative agencies serving as adjuncts to Article III courts—as would be the case with the proposed FCA courts since step five involves judicial review in a district court—may make findings of fact subject only to a higher standard of review. But, findings of law must face de novo review in an Article III court. Within the existing Medicare system, into which the federal healthcare courts could integrate, the Medicare Appeals Council’s legal conclusions are reviewable de novo, and findings of fact are subject to substantial evidence review. Even though questions of law will be subject to de novo review, the multiple layers of expert review by specialized courts with their own appellate panels will lend greater credence to the opinions, thus making the Article III courts hesitate before reversing. Consequently, situations like the outcome in the FCA “split”—such as where the Third Circuit completely rejected the district court’s analysis—would be avoidable. Additionally, courts give

---

181 The district court ruled that:
The difference of opinion of an expert cannot be false . . . . diverging opinions do not create a genuine issue of material fact about the falsity of a physician's determinations that the patient meets hospice eligibility where, as here, there is no factual evidence that Defendant's certifying doctor was making a knowingly false determination. This is because the ultimate issue is not whether the certification of hospice eligibility was correct or incorrect, but rather whether it was knowingly false.


182 See United States ex rel. Druding v. Druding, 952 F.3d 89, 100 (3d Cir. 2020).


184 Id.


186 See Howard & Maine, supra note 166 and accompanying text.
substantial deference to the agency’s reasonable interpretations, even when conducting a de novo review. Further, the specialty court opinions will set forth factual findings that will benefit from a substantial evidence standard, which will help address the complexities of the underlying medicine and free up the Article III courts to focus on legal issues when reviewing appeals. This will insulate the medical facts from non-specialized Article III judges lacking medical backgrounds.

b. Specialty Healthcare Courts are Constitutional

Before explaining how these courts will solve the problem demonstrated by the FCA confusion, it is important to address the threshold issue of whether such courts are constitutional. Article III of the U.S. Constitution established the Supreme Court and gave Congress the power to create lower Article III courts to preside over the types of cases enumerated therein. Article III judges benefit from life tenure, assuming good behavior, as well as salaries that cannot be decreased during the judges’ terms of office. Article III grants jurisdiction over various enumerated cases and controversies. Applied to the FCA, which Congress enacted in 1863, Article III courts have

---

190 See U.S. CONST. art. III, § 1 (“The judicial Power of the United States shall be vested in one supreme Court, and in such inferior Courts as the Congress may from time to time ordain and establish.”).
191 See id.
192 Article III of the Constitution states:

The judicial Power shall extend to all Cases, in Law and Equity, arising under this Constitution, the Laws of the United States, and Treaties made, or which shall be made, under their Authority;—to all Cases affecting Ambassadors, other public Ministers and Consuls;—to all Cases of admiralty and maritime Jurisdiction;—to Controversies to which the United States shall be a Party;—to Controversies between two or more States;—between a State and Citizens of another State;—between Citizens of different States;—between Citizens of the same State claiming Lands under Grants of different States, and between a State, or the Citizens thereof;—and foreign States, Citizens or Subjects.

See U.S. CONST. art. III, § 2.
jurisdiction under the federal question doctrine.\textsuperscript{194} Other Medicare and Medicaid lawsuits addressed by existing specialty courts also involve federal questions because they likewise deal with federal statutes.\textsuperscript{195}

The Constitution empowers Congress to create Article III specialized courts.\textsuperscript{196} For example, the U.S. Court of International Trade is an Article III court with “nationwide jurisdiction over civil actions arising out of the customs and international trade laws of the United States.”\textsuperscript{197} Congress could similarly create an Article III court with jurisdiction over civil actions arising from federal healthcare laws, such as the FCA, Medicare, and Medicaid. If Congress did this, no constitutional issues would arise, provided judges have life tenure and salary protection.\textsuperscript{198}

More often, Congress creates specialty courts under Article I, (sometimes referred to as legislative courts) to handle complex areas of law.\textsuperscript{199} For example, bankruptcy courts are non-Article III courts,\textsuperscript{200} and the Environmental Protection Agency, Social Security Administration, and Employee Benefits Security Administration, also not created under Article III, all make use of administrative law judges.\textsuperscript{201} These judges have the requisite

\textsuperscript{194} See U.S. Const. art. III, § 2.
\textsuperscript{195} Medicare is a federal statute, and Medicaid is a federally funded program. See 42 U.S.C. § 1396 (2022); Financial Management, supra note 46.
\textsuperscript{198} See Congressional Power to Establish Article III Courts, supra note 196.
expertise to address the complicated issues involved in the relevant practice areas. For instance, merit selection panels, which are often largely composed of bankruptcy practitioners, choose bankruptcy judges.202 While there is no requirement that new judges possess bankruptcy experience, the bankruptcy community is very exclusive.203 Indeed, many judges obtain their positions after hearing about vacancies through word-of-mouth or personal relationships in the bankruptcy community.204

Legal scholars have debated whether the Constitution authorizes Congress to create non-Article III courts.205 The constitutional objection to non-Article III courts is that Congress might weaken the judicial branch by removing some of its power and reallocating it to judges lacking the independence of Article III. Specifically:

Article I contains no guarantee that the judges of Article I courts have life appointments. Nor does it provide that their salaries may not be reduced during their term of office. On the other hand, the tenure of an Article III judge is during “good behaviour”; moreover, Article III provides that its judges shall have a compensation that “shall not be diminished during their Continuance in Office.”206

Nonetheless, for 200 years, Congress has created courts without the tenure and salary protections of Article III and given them...
authority to adjudicate Article III matters, a factor weighing in favor of their constitutionality.

The generally accepted circumstances include three “narrow exceptions” to Article III: territorial courts, military courts, and the adjudication of “public rights.” Public rights are defined as “disputes between the Government and others,” not including criminal matters. More recently, the Court has allowed non-Article III courts that might not fall into one of those three exceptions so long as “the essential attributes of judicial power are retained in the art. III court.” As the Court has explained:

Congress possesses the authority to assign certain factfinding functions to adjunct tribunals. It is, of course, true that while the power to adjudicate “private rights” must be vested in an Art. III court, this Court has accepted factfinding by an administrative agency, as an adjunct to the Art. III court, analogizing the agency to a jury or a special master and permitting it in admiralty cases to perform the function of the special master.

“Private rights” address “private unalienable rights of each individual,” such as one individual’s liability to another, and are inherently judicial. This is contrasted with “public rights” that

---

207 See CHEMERINSKY, supra note 205, at 234; see, e.g., I.R.C. §§ 7441, 7446 (1982) (creating Tax Court, where judges sit for fifteen-year terms); Atlas Roofing Co. v. Occupational Safety & Health Review Comm’n, 430 U.S. 442, 460–61 (1977) (discussing the constitutionality of Congress empowering the Occupational Safety and Health Commission, an administrative agency, to impose civil penalties for matters within the cases and controversies enumerated in Article III).

208 See CHEMERINSKY, supra note 205, at 235–36 (citing American Ins. Co. v. Canter, 26 U.S. (1 Pet.) 511 (1828) (noting the Supreme Court has long recognized the constitutionality of non-Article III courts).

209 See N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 458 U.S. 50, 64–68, 94 (1982) (plurality opinion), superseded by statute, Bankruptcy Amendments and Federal Judgeship Act of 1984, Pub. L. No. 98-353, 98 Stat. 333, as recognized in Wellness Int’l Network, Ltd. v. Sharif, 575 U.S. 665 (2015). Public rights generally refer to cases where private citizens sue the government; however, non-Article III courts and administrative agencies are often granted authority under the public rights doctrine to assess penalties on private individuals, despite the lack of life tenure for administrative law judges and commissioners. See CHEMERINSKY, supra note 205, at 237. Indeed, the Supreme Court acknowledged “[f]amiliar illustrations of administrative agencies created for the determination of [public rights] matters are found in connection with the exercise of the congressional power as to public health.” Crowell v. Benson, 285 U.S. 22, 51 (1932). Therefore, while this Note proceeds under the adjunct exception leaving the “essential attributes of judicial power” to Article III courts, it is worth noting there might also be a public rights argument justifying the creation of federal healthcare courts. N. Pipeline Constr. Co., 458 U.S. at 81.

210 See id. at 69–70, n.24.

211 Id. at 81.


are not inherently judicial because they can start in the courts but can also be resolved by the executive and legislative branches.\textsuperscript{215}

The proposed federal healthcare courts fit within this constitutional framework. The Supreme Court treats federal statutes involving quasi-public rights akin to public rights, condoning review by non-Article III courts without consent of the parties and with little review.\textsuperscript{216} Specifically, in connection with the Federal Insecticide, Fungicide, and Rodenticide Act provision authorizing the Environmental Protection Agency to consider data already in its files when evaluating a new applicant’s request for “if the applicant has made an offer to compensate the original data submitter,” the Supreme Court addressed the constitutionality of a federal law mandating binding arbitration with limited judicial review for resolving disputes among private parties that fail to agree on a compensation amount.\textsuperscript{217} It upheld the constitutionality of the arbitration provision, finding that “Congress, acting for a valid legislative purpose pursuant to its constitutional powers under Article I, may create a seemingly ‘private’ right that is so closely integrated into a public regulatory scheme as to be a matter appropriate for agency resolution with limited involvement by the Article III judiciary.”\textsuperscript{218} Specifically, the private right to compensation in \textit{Thomas} was integral to the federal regulatory scheme of encouraging competition and streamlining research, because it spread the cost among applicants instead of each applicant repetitively shouldering the entire cost individually.\textsuperscript{219} Similarly, the existence of compensation for relators in \textit{qui tam} causes of action is integral to the federal scheme of rooting out fraud because it encourages individuals to assist the government with enforcement by bearing the burden of the cost and time investments associated with prosecution.\textsuperscript{220} Indeed, legal scholars classify \textit{qui}

\textsuperscript{217} See \textit{id. at} 571, 573–74.
\textsuperscript{218} \textit{Id. at} 593–94; see also \textit{Commodity Futures Trading Comm’n v. Schor}, 478 U.S. 833, 835–36, 858 (1986) (finding even ancillary jurisdiction of state law counterclaims constitutional where the Commodity Futures Trading Commission (“CFTC”) adjudicated “reparations procedure through which disgruntled customers of professional commodity brokers could seek redress for the brokers’ violations of the [Commodity Exchange] Act or CFTC regulations”).
\textsuperscript{219} See \textit{Thomas}, 473 U.S. at 570.
tam actions, like the FCA, as quasi-public rights. Accordingly, FCA lawsuits, where the government leaves the litigation in the hands of relators who share in a portion of the recovery, similarly involve a right to compensation under federal law closely related to a public regulatory scheme. Thus, non-Article III adjudication for those cases should likewise be deemed constitutional.

Even if FCA claims are not quasi-public when involving government-initiated civil litigation—and therefore “inherently judicial”—use of a non-Article III adjunct would still be appropriate because the healthcare courts’ power is limited and there is adequate review in an Article III court. In Crowell v. Benson, the Supreme Court upheld a requirement that workers injured in maritime accidents file their claims with the U.S. Employees’ Compensation Commission. The Court reasoned the Commission was constitutional because it functioned as an adjunct to Article III courts. Specifically, the Commission lacked independent authority to enforce compensation orders, which were instead appealable to federal district courts, and Article III courts possessed de novo review of questions of law, constitutional facts, and jurisdictional facts.

The Commodity Futures Trading Commission employs this same appeal structure for reparations it orders for individuals injured by brokers’ fraudulent or illegally manipulative conduct. In Schor, the Court found the Commission’s exercise of this power to be “of unquestioned constitutional validity.” The real constitutional entanglement emerged in addressing the Commission’s power to adjudicate counterclaims arising from the

221 In clarifying the distinction between private and quasi-public rights, Justice Thomas relied on a law review comment that classified the individual’s right to bring qui-tam actions as a quasi-private “privilege”] that the government could validly supplant any time before judgment. See Teva Pharm. USA, Inc. v. Sandoz, Inc., 574 U.S. 318, 344 n.2 (2015) (Thomas, J., dissenting); Caleb Nelson, Adjudication in the Political Branches, 107 COLUM. L. REV. 559, 571 (Apr. 2007). Per Justice Thomas, “no matter how closely a franchise resembles some ‘core’ private right, it does not follow that it must be subject to the same rules of judicial interception as its counterpart.” Teva Pharm., 574 U.S. at 344 n.2.


225 See id. at 41, 53–54.


227 See Schor, 478 U.S. at 856.
same conduct, because this went beyond the traditional agency model.228 Here, the Court leaned heavily on the idea of consent to uphold the Commission’s constitutional validity.229

Adjuncts also adjudicate bankruptcy cases, which likewise involve private rights, with consent.230 Initially, the Supreme Court found the grant of jurisdiction to bankruptcy courts unconstitutional,231 and issued a plurality opinion stating bankruptcy courts could not be considered adjuncts to Article III courts because their jurisdiction was not limited to a specific area of law, but extended to all civil matters.232 A concurring opinion that struggled with the bankruptcy court’s authority to adjudicate state law matters only loosely related to bankruptcy law.233 Of note, neither of these constitutional concerns would pose a problem for the proposed federal healthcare courts, which would have jurisdiction over a specific area (healthcare) and would not entangle with state law matters. However, bankruptcy courts are of course still operating today, with the option for parties to appeal to the Bankruptcy Appellate Panel (“BAP”).234 The constitutional defects were remedied by the limit of jurisdiction to “core” proceedings involving debtor’s property, whereas “noncore” matters cannot be heard by the bankruptcy courts, except the issuance of proposed findings of fact and law for noncore matters with an independent basis for federal jurisdiction.235

Similar to the commissions in Crowell and Schor, the proposed federal healthcare courts opinions would address fraud, among other healthcare statutes, and could ultimately be appealed to district courts, where legal conclusions therein would face de novo review. While Schor leaned on the idea of consent to uphold, and bankruptcy courts had to limit the review of “noncore” matters and rely on a consent model for BAP, federal healthcare courts do not need to incorporate consent because they do not pose the same constitutional concerns. Even with a mandatory structure that has a specialized appeals process through the Medicare Appeals Council (akin to BAP), the
ultimate decision-making remains with the independent Article III judiciary via the final step in the appeals process. Indeed, making the healthcare courts hinge on the parties’ consent would undermine the goals of federal healthcare courts to provide multiple layers of guidance to unspecialized judges as a way of insulating the medical component of the rulings from misinterpretation. In short, health care courts are constitutional as public rights courts because, as structured, they will leave the “essential attributes of judicial power” with Article III courts.

C. Benefits of Health Courts

1. Expertise Would Ameliorate Accuracy and Efficiency Concerns

Expanding the Medicare adjudication system into broader health courts would address the problems of inaccurate rulings and clogged courts because these health courts would employ specialized health care judges. As a function of these judges developing a significant level of expertise in constantly overseeing healthcare lawsuits, they would be expected to become excellent fact finders which would promote improved quality in rulings. Specific to the FCA “split,” the Third Circuit’s ruling in Druding mischaracterized the Eleventh Circuit’s ruling in AseraCare because it hyper-focused on analyzing “objective falsity” and its scienter element. A healthcare judge with a better understanding of medicine would have been able to successfully parse the Eleventh Circuit’s application of the law to the facts, realize hospice certifications are complex and account for numerous imprecise factors, and discern that an expert physician disagreeing with the treating physician’s conclusion does not indicate the treating physician’s conclusion was false, or even erroneous. Accordingly, a healthcare judge would not have overlooked the reasonableness standard, and having noticed that such legal standard proved key to the case, would not have wasted time and resources on the confusing and irrelevant scienter discussion in Druding.

---

238 See id. at 639.
239 See United States ex rel. Druding v. Druding, 952 F.3d 89, 96 (3d Cir. 2020).
240 See supra notes 53–65 and accompanying text.
241 See United States v. AseraCare, Inc., 938 F.3d 1278, 1287 (11th Cir. 2019) (noting the expert physician clarified his review was based on “his own clinical judgment”).
Moreover, specialized judges with a deeper understanding of medicine would not need to spend as much time familiarizing themselves with the medicine for each case because they would already have a strong baseline. Evidence of this efficiency is demonstrated by specialty courts adjudicating matters more quickly than traditional courts. For example, bankruptcy is very similar to medicine in that it also involves its own sort of language, and without an understanding of the bankruptcy jargon, a judge cannot hope to adjudicate bankruptcy matters properly. Bankruptcy Appellate Panels have demonstrated an ability to ease the burden on the docket with faster disposition as well as fewer appeals than their district court counterparts. Bankruptcy Appellate Panels have an average resolution timeframe of 8.6 months with many cases handled in even shorter periods of time as procedural issues are resolved. Given the parallel of complex terminology, logically health courts would accelerate judicial resolution of healthcare lawsuits much in the same way as bankruptcy courts.

2. The FCA (and Other Federal Healthcare Litigation) Constitute a Large Enough Portion of Government Revenue to Financially Justify the Recommended Health Courts

Although establishing these health courts could constitute a big undertaking, it is a well-justified cost that is lessened by piggybacking off the existing Medicare Appeals Council. The creation of healthcare courts would not only serve as a venue for FCA cases but would also serve to adjudicate other healthcare matters, including Medicare and Medicaid cases. Moreover, in addition to the FCA, multiple statutes govern Medicare fraud and abuse including the Physician Self-Referral Law (“Stark Law”) and Civil Monetary Penalties Law (“CMPL”).


244 See U.S. Cts., supra note 242.

245 See id.

246 See generally County of Los Angeles v. Shalala, 192 F.3d 1005, 1008 (D.C. Cir. 1999) (interpreting requirements of the Medicare statute); Orthopaedic Hosp. v. Belshe, 103 F.3d 1491, 1492 (9th Cir. 1997) (considering whether Medi-Cal hospital outpatient rates violated the federal Medicaid Act).

Healthcare-related fraud, including that involving hospice organizations, laboratories, medical device manufacturers, drug companies, pharmacies, managed care providers, hospitals, and physicians, accounts for more than $5 billion of the $5.6 billion in total FCA settlements and judgments. Healthcare fraud settlements and judgments primarily focus on Medicare, Medicaid, and TRICARE, which serves the military. Not included in the data are savings realized as a consequence of deterring fraud via vigorous prosecution.

For fiscal year 2021, 701 new FCA-related matters were filed, including 203 non qui tam and 598 qui tam cases with settlements and judgments totaling $3,984,299,554. Of this total, the Department of Health and Human Services was responsible for $3,590,882,626, broken down into 97 non qui tam and 388 qui tam cases. In 2020, despite a pandemic, the 934 new FCA cases filed represented the largest single year total, correlating to a significant percentage of the 4,125 new cases over the last five years. Of note, healthcare recoveries represent over eighty percent of the past five years’ worth of recoveries.

The government has also accelerated its involvement in rooting out fraud on its own without whistleblowers via various types of data analysis used to identify patterns of excessive billing to government programs which are then flagged for potential fraud.
The federal government has historically recognized healthcare fraud as a priority, and establishing specialized health courts would be consistent with this goal. The Senate and House of Representatives have held hearings dedicated entirely to fighting healthcare fraud. More recently, in a February 2021 speech, Acting Assistant Attorney General Brian Boynton detailed the priorities of FCA enforcement. Those priorities are pandemic-related fraud, opioids, fraud targeting seniors, electronic health records, telehealth, and cybersecurity. Each area discussed related in some fashion to healthcare, making healthcare fraud the Civil Division’s clear-cut current prosecutorial objective. In connecting healthcare issues to each category, Boynton referenced pandemic-related healthcare concerns, elderly patients receiving poor or unnecessary healthcare, and the risk of cyberattacks targeting government data including medical records.

3. Expanding the Existing System to Create Health Courts Would Financially Benefit Consumers

When insurers are forced to pay out claims, they reallocate those costs by increasing premiums and deductibles for policyholders. When policyholders are service providers, such as hospitals, they raise the cost of services to counteract increased liability expenses. The increased cost of medical services is next shifted to the patient’s health insurance company, which in turn


257 See Assistant A.G. Boynton Remarks, supra note 255.

258 See id.

259 See id.

260 See id.


raises premiums and deductibles for policyholders and decreases coverage. Either the patient directly bears the burden of these costs as the policyholder forced to pay increased premiums and deductibles for less healthcare coverage, or the patient shares this burden with his employer. When companies pay these increased premiums for their employees as part of a health insurance benefit program, the burden can ultimately land on consumers due to the increased cost of doing business, or come out of the employee’s compensation.

Respecting the FCA, this cost shifting affects patients in two ways. First, corporations have been assessed billions of dollars in penalties stemming from FCA violations over the past decade, generating claim payments through professional liability insurance policies, with numbers of policy holders seeking coverage continuing to increase. The sheer volume of recoveries, exceeding $22 billion to companies over the last six years, clearly has a significant impact on both underwriting and claims assessments.

Second, according to the National Health Care Anti-Fraud Association, healthcare fraud costs the United States tens of billions of dollars annually, accounting for at least three percent of total expenditures, while others claim this figure could run as high as ten percent. The Federal Bureau of Investigation reports that fraudulent billing constitutes the most serious of

---

263 See Allen, supra note 261.
265 See Health Insurance Coverage of the Total Population, KAISER FAM. FOUND., http://www.kff.org/other/state-indicator/total-population/?dataView=0&currentTimeframe=0&sortModel=%7B%22colId%22:%22%22008__Non-Group%22,%22sort%22:%22%22desc%22%7D [http://perma.cc/C83H-EGU3] (last visited Dec. 20, 2022) (acknowledging in 2019, 5.9% of Americans purchase health insurance policies directly from the insurer, instead of through an employer).
267 See The Challenge of Health Care Fraud, supra note 264; see also Allen, supra note 261.
269 See id.
270 See The Challenge of Health Care Fraud, supra note 264.
these offenses.\textsuperscript{271} This translates directly into consumer losses because, as discussed above, increasing the amount billed to patients’ health insurance companies results in increased insurance premiums and coverage limits.\textsuperscript{272}

These two problems might seem at odds with each other—the first seemingly advocating to lower FCA liability, and the latter to increase it. Nonetheless, this tension can be reconciled. Accuracy is key. Going too far would over-impose liability and result in excessive professional liability insurance payouts, where cost shifting would ultimately place the burden on consumers.\textsuperscript{273} Not doing enough will result in a lax system ineffective at rooting out and deterring fraud—fraud that may ultimately take money out of consumers’ pockets.\textsuperscript{274} While our current court systems are not achieving this needed accuracy when it comes to healthcare rulings,\textsuperscript{275} specialized health courts could.\textsuperscript{276}

CONCLUSION

The Third Circuit’s perception of a deep circuit split in physician liability under the FCA, and its corresponding erroneous representation of the existing caselaw, demonstrates the need for specialized review of all federal healthcare cases, including FCA issues, and not just those already addressed in the Medicare appeals system. This need for reform is further demonstrated by the clogged court system which could be relieved by increased efficiency of expert judges, the medical community’s history of advocating for health courts due to discontent with the current system, and the financial considerations at play on the government and consumer level.

Congress should answer this call to action by creating an administrative agency as an expansion upon the existing Medicare Appeals Council to handle all civil federal healthcare cases. The healthcare expert judges employed by these courts would issue more accurate rulings and remedy efficiency concerns, ultimately benefiting medical providers, patients, and the government alike.


\textsuperscript{272} See \textit{The Challenge of Health Care Fraud}, supra note 264.

\textsuperscript{273} See supra notes 261–268 and accompanying text.

\textsuperscript{274} See supra notes 266–272 and accompanying text.

\textsuperscript{275} See discussion supra Part I.C.

\textsuperscript{276} See discussion supra Part II.C(i).
APPENDIX I: FCA CASES AND RECOVERIES ATTRIBUTABLE TO HEALTHCARE

The statistics below were obtained from the Civil Division, U.S. Department of Justice’s Fraud Statistics – Overview: October 1, 1986 - September 30, 2021, and Fraud Statistics – Health and Human Services: October 1, 1986 - September 30, 2021.

Stars on the Field, Benchwarmers on the Tax Return: Student-Athletes and the Tax Ramifications of Name, Image, and Likeness Deals

Haley A. Ritter

“The hardest thing in the world to understand is the income tax.”

- Albert Einstein

INTRODUCTION.................................................................402

I. A BRIEF BACKGROUND OF COLLEGE ATHLETICS AND THE RESTRICTIONS ON PAYMENTS TO STUDENT-ATHLETES 404
   A. History of College Athletics and the Pay-for-Play Debate ............................................................... 404
   B. Challenges to the Traditional Reluctance to Allow Student-Athletes to Receive Pay .......... 408
      1. State Legislation ................................................. 408
      2. NCAA v. Alston ............................................. 410
   C. Further Changes to Paying Student-Athletes since NCAA v. Alston ............................... 412
      1. NCAA Interim NIL Policy ............................. 412
      2. NIL Legislation ............................................. 413

II. FEDERAL TAXATION OF STUDENT-ATHLETES ............. 416
   A. Evolution of Taxation of Student-Athletes and the Decrease of Preferential Treatment .......... 418
   B. Complications with Independent Contractor Status ................................................................. 420

* J.D. Candidate, Expected May 2023, Chapman University Dale E. Fowler School of Law. B.S. Tulane University, 2018. I am grateful to Bobby Dexter for guidance and invaluable feedback during the writing process and my fellow editors on Chapman Law Review for thorough revisions. I am also thankful to Christian Ritter for support and encouragement through this project.

INTRODUCTION

Ben has trained his whole life for this moment. The January sun is glaring down on the Indianapolis football field for the 2023 College Football Playoff National Championship. There are 24 seconds left on the clock, and Ben’s team is five points behind. They have worked so hard to get to this point in the season, and they cannot give it up now. The center snaps the ball to Ben, who drops back into the pocket while surveying the field, frantically searching for an open receiver. Everyone is covered, so he is forced to roll out of the pocket and make the play with his legs. As he nears the line of scrimmage, a wide receiver breaks away from the defensive back. With no time to think, Ben lays the ball out deep down the field. Miraculously, the wide receiver hauls in the pass, scoring the game-winning touchdown with only 6 seconds left on the clock. Ben simply cannot believe it; all his dreams are coming true! He and his teammates are National Champions! The crowd is roaring, the team is ecstatic, and Ben is having the best day of his life!
Over the next few days, Ben sees his face all over Adidas's Instagram. He had completed an exhausting photoshoot for the company before the championship game, so there was plenty of material to work with pursuant to his name, image, and likeness ("NIL") deal. Ben’s contract with Adidas was the biggest NIL deal to date for a college player. It was a four-year commitment which, barring material adverse events, promised him a minimum of eight million dollars over the deal term. But the benefits certainly did not come without costs, burdens, and risks. Ben has never had to file a tax return before. In fact, he has never even seen an IRS Form 1040. However, now he has the overwhelming burden of submitting both federal and state tax returns (having no idea what proper tax compliance entails), while also prioritizing academics, attending football practice, performing well at games, and holding up his end of a multi-million dollar deal with Adidas.

Ben’s financially gratifying situation is now the sobering reality for many student-athletes at colleges around the country. Recent common law and statutory changes have paved the way for student-athletes to earn compensation from their NIL deals. Such revolutionary progress regarding student-athlete rights does not, however, come without ramifications. Taxes are a necessary evil. Everyone must pay their fair share, but filing timely and accurate tax returns can be difficult. Even if a student-athlete (fresh out of high school) has managed to secure a lucrative NIL deal, they may have never faced the daunting challenge of filing a complicated tax return. Moreover, such individuals have both federal and state-level tax hurdles to clear.

This Article addresses the increased complexity and heightened tax-compliance requirements stemming from student-athletes’ newfound freedom to receive education-related benefits from universities and NIL payments from third parties. These student-athletes are too young and inexperienced to independently maneuver the intricacies surrounding these tax implications. As such, a solution must be found to prevent the student-athletes from unnecessary burdens and unforeseen

\[\text{See, e.g., Nat’l Collegiate Athletic Ass’n v. Alston, 141 S. Ct. 2141 (2021) (Kavanaugh, J., concurring) (discussing how the NCAA’s limitation of student-athletes’ income is generally a violation of antitrust law); S. 206, 2019-2020 Leg., Reg. Sess. (Cal. 2019).}\]

\[\text{See infra Parts III & IV.}\]

\[\text{See Compania Gen. De Tabacos De Filipinas v. Collector of Internal Revenue, 275 U.S. 87, 100 (1927) (Holmes, J., dissenting) (“Taxes are what we pay for civilized society.”); see also New York ex rel. Cohn v. Graves, 300 U.S. 308, 313 (1937) (“Enjoyment of the privileges of residence in the state and the attendant right to invoke the protection of its laws are inseparable from responsibility for sharing the costs of government.”).}\]

\[\text{See Lipman, supra note 1, at 1.}\]
penalties. Part I discusses the background of the pay-for-play debate in college athletics and how the NCAA v. Alston decision changed the face of college sports. Part II walks through important aspects of federal taxes that student-athletes must consider in light of their newfound financial freedom stemming from the NCAA v. Alston decision. Part III follows with provides a similar discussion regarding state taxes. The federal and state obligations of student-athletes, who are young, inexperienced taxpayers, will be highly burdensome and immensely complicated. Part IV sets forth recommendations to help student-athletes navigate the tax ramifications of their income stemming from NIL deals.

I. A BRIEF BACKGROUND OF COLLEGE ATHLETICS AND THE RESTRICTIONS ON PAYMENTS TO STUDENT-ATHLETES

A. History of College Athletics and the Pay-for-Play Debate

Before intercollegiate sports existed, college students would participate in elaborate and violent intramural contests, called “class rushes.” These informal contests eventually gained some organization, and students started competing against other schools. The first intercollegiate competition was a boat race between Harvard and Yale in 1852. The event was sponsored by a railroad superintendent who provided the competitors “an all-expenses-paid vacation with lavish prizes—along with unlimited alcohol.” College sports at this time were still unregulated and

---

6 Alston, 141 S. Ct. at 2141.
7 See infra Part I.
8 See infra Part II.
9 See infra Part III.
10 See infra Parts II–III.
11 See infra Part IV.
13 See id.
extremely violent. For example, the rules for football allowed anything. Offensive players could pick up and throw other players into the endzone to score a touchdown, and defensive players could throw their backs at the offensive players to prevent them from scoring. Because of the lack of rules and protective equipment, between 1880 and 1905, football caused 330 deaths and 1,149 serious injuries. These statistics reflect an average of 13 deaths and 45 serious injuries per year.

By 1896, a few midwestern universities formed the Western Conference (now known as the Big Ten Conference). By 1905, as a direct response to the immense dangers of college football, President Theodore Roosevelt summoned thirteen coaches and administrators and gave them a choice to either “[r]eform football’s rules or abolish the game.” Despite President Roosevelt’s ultimatum, they failed to act, and six weeks later, another college football fatality occurred. This incident motivated the coaches and college administrators to take action, and they created the Intercollegiate Athletic Association of the United States (“IAAUS”) to regulate college athletics. By its second meeting in 1906, the IAAUS created rules that revolutionized college football. The IAAUS “adopted the forward pass, established the one-yard neutral zone, cut game time from 70 to 60 minutes, required six men on the offensive line, and mandated a 10-yard gain (instead of the previous five) for a first down.” The original by-laws included a provision recommending that “[n]o student shall represent a College or University in any intercollegiate game or contest who is paid or

17 Id.
18 Id.
19 Id.
21 See Rushin, supra note 16.
22 See id.
23 See id.
24 See id.
25 Id.
receives . . . any money.” By 1911, the IAAUS changed its name to the National Collegiate Athletic Association (“NCAA”).

Despite the NCAA’s recommendation against compensating student-athletes, many colleges were still using money to induce players to join their collegiate sporting programs. As more conferences began to develop, the Southeastern Conference (“SEC”) and the Big Ten Conference (“Big Ten”) started to compete for national dominance. The SEC was recruiting student-athletes by promising them large monetary benefits. The Big Ten attempted to end the bidding war on student-athletes and maintain a competitive balance among universities. To accomplish this goal the Big Ten members lobbied college coaches and athletic directors, arguing that they should expand the NCAA’s power and protect the principle of amateurism within college football. Simultaneously, concerns arose that these student-athletes could be viewed as employees, subjecting universities to labor rules concerning wages, overtime, and workers’ compensation. As a result, the NCAA promulgated “The Principle of Amateurism.” According to this principle, “[s]tudent-athletes shall be amateurs in an intercollegiate sport, and their participation should be motivated primarily by education and by the physical, mental and social benefits to be derived.”

Today, the NCAA is made up of 1,117 schools which are organized into 40 conferences and 3 divisions. Division I is comprised of the most competitive and rigorous college athletic programs. These Division I athletics programs traditionally have higher budgets than other programs. Within Division I, there are

---

28 See Alston, 141 S. Ct. at 2149; see also Kelly Charles Crabb, The Amateurism Myth: A Case for a New Tradition, 28 STAN. L. & POL’Y REV. 181, 190 (2017). Student-athletes were being paid so much that, in the 1940s, Hugh McElhenny took a cut in salary when he switched from college to professional football. See Alston, 141 S. Ct. at 2149.
29 See Crabb, supra note 28.
30 See id.
31 See id. at 190, 193.
32 See id. at 191.
34 See NCAA DIVISION I MANUAL, supra note 33.
35 See Rollins, supra note 27.
37 Id.
subdivisions, with the Football Bowl Subdivision (“FBS”) being the highest caliber and most prominent subdivision.38

Historically, the NCAA has created a reputation of “play[ing] a critical role in the maintenance of a revered tradition of amateurism in college sports.”39 Student-athletes lose their eligibility to compete in college athletics governed by the NCAA if they lose their amateur status.40 Students have lost their eligibility by making videos that were reposted on ESPN and by making money on videos they posted on YouTube.41 Injured players have been unable to obtain worker’s compensation because they were not employees of the university when they were injured.42

Per the NCAA Division I Manual, student-athletes could earn compensation for actual work performed at a job if that pay was commensurable to the usual pay rate for the job performed, but they could not use their name, image, and likeness for promotion.43 The only money a college or university may give to student-athletes must be in the form of a scholarship.44 These scholarships must be limited to the cost of tuition, required institutional fees, the cost of room and board, course-related books, and other expenses related to attendance, up to the cost of attendance.45 The

---

40 See NCAA DIVISION I MANUAL, supra note 33, at 63 (“An individual loses amateur status and thus shall not be eligible for intercollegiate competition in a particular sport if the individual . . . [u]ses athletics skill (directly or indirectly) for pay in any form in that sport . . . .”).
42 But cf. @InsidetheNCAA, TWITTER (July 31, 2017, 1:40 PM), http://twitter.com/InsidetheNCAA/status/892122868355657728 [http://perma.cc/FS8M-Q9D5] (claiming that De La Haye would not have lost his NCAA eligibility if he had separated athletically-related videos from non-athletic videos and only monetized on his non-athletic videos).
44 See NCAA DIVISION I MANUAL, supra note 33, at 193.
45 See id. at 189–91.
only third parties from which student-athletes may receive money are their parents/guardians and certain financial aid programs. The NCAA has traditionally opined against and prohibited payments to student-athletes for their NIL as a part of their commitment to the principle of amateurism. In general, the NCAA has traditionally set strict limitations on any money that student-athletes may receive and refused to allow a pay-for-play model within college athletics.

B. Challenges to the Traditional Reluctance to Allow Student-Athletes to Receive Pay

Recently, scholars have argued that student-athletes are employees, regardless of the language used by the NCAA. Furthermore, courts and legislatures have begun reinstating student-athletes' rights to receive compensation. Even the NCAA has “admitted that restrictions on student-athlete compensation should be loosened or eradicated.”

1. State Legislation

California became the first state to enact legislation allowing student-athletes to profit from their NIL when it passed the Senate Bill, “Collegiate athletics: student athlete compensation and representation,” commonly referred to as the Fair Pay to Play Act (“FPTPA”). The FPTPA prohibited colleges, universities, athletics associations, conferences, and the NCAA from preventing student-athletes from earning compensation from their NIL. The FPTPA’s text as originally introduced made clear that profiting off

46 See id. at 192.
47 See Nat’l Collegiate Athletic Ass’n v. Alston, 141 S. Ct. 2141, 2151–52 (2021) (“The NCAA’s only remaining defense was that its rules preserve amateurism, which in turn widens consumer choice by providing a unique product—amateur college sports as distinct from professional sports.”) But see Steve Berkowitz, Oliver Luck Brings Own Perspective to NCAA on O’Bannon Name and Likeness Issue, USA TODAY SPORTS (Jan. 16, 2015, 6:05 PM), http://www.usatoday.com/story/sports/college/2015/01/16/ncaa-convention-oliver-luck-obannon-name-and-likeness-court-case/21873331/ [http://perma.cc/8YKW-KNUJ] (noting that NCAA executive Oliver Luck’s view in support of NIL payments is contrary to that of the NCAA).
52 See S. 206, 2019-2020 Leg., Reg. Sess. (Cal. 2019); see also Paul McDonnell, California’s Fair Pay to Play Act and its Fight Against the Commerce Clause, 39 J.L. & COM. 75, 75–76 (2020) (“California Senate Bill 206, more aptly referred to as the Fair Pay to Play Act, has sent shockwaves through the intercollegiate athletic community.”).
their NIL “shall not affect [a] student’s scholarship eligibility.” 54

This originally introduced text of the FPTPA also included some statistics from a study explaining why student-athletes should be able to receive money for their NIL. 55 For example, eighty-two percent of student-athletes living on campus and ninety percent of student-athletes living off campus were living below the poverty level during the 2010-2011 academic year. 56 In comparison, the coaches for those student-athletes made $3,500,000 that same year. 57 The Act also included an estimation from a study finding that from 2011 to 2015, Division I FBS men’s football and basketball players forfeited $6,200,000,000 of potential earnings. 58 California has 24 of the 254 Division I football teams, including 7 which are part of the Division I FBS. 59 Therefore, a large portion of the harm to student-athletes—who miss out on profiting from their NIL—occurs in California. 60

The NCAA responded to the FPTPA by claiming that it would “erase the critical distinction between college and professional athletics” and upend the balance of maintaining an even playing

---

56 See S. 206, 2019-2020 Leg., Reg. Sess. § 1(a)(2) (Cal. 2019) (as introduced by Senate, Feb. 4, 2019). The coaches are not the only people profiting off the student-athletes:
Those who run this enterprise profit in a different way than the student-athletes whose activities they oversee. The president of the NCAA earns nearly $4 million per year. Commissioners of the top conferences take home between $2 to $5 million. College athletic directors average more than $1 million annually. And annual salaries for top Division I college football coaches approach $11 million, with some of their assistants making more than $2.5 million.
field across the nation.61 The NCAA clarified that it was working on creating a fair way for student-athletes to use their NIL to make income, but that it would not allow for a pay-for-play model, reiterating that “[t]he NCAA has consistently stood by its belief that student-athletes are students first, and they should not be employees of the university.”62 The NCAA urged California to reconsider the “harmful” and “unconstitutional” bill.63

After California’s FPTPA was enacted, many states followed closely behind.64 As of March 2021, thirty-five states had introduced NIL legislation, six of which had already been passed.65

2. NCAA v. Alston

On June 21, 2021, a case was heard by the Supreme Court of the United States that changed the game of paying student-athletes. In NCAA v. Alston, current and former Division I FBS football and men’s and women’s Division I basketball student-athletes asserted a class action against the NCAA and eleven Division I Conferences, claiming that the NCAA’s limits on student-athletes’ compensation violated federal antitrust law.66 The Alston class argued that the NCAA held a monopsony over intercollegiate athletics because it used its power to restrict student-athletes’ pay below the market level.67 In response, the NCAA argued that limiting student-athlete income preserved the brand of amateurism upon which the NCAA was built.68 The district court held that the NCAA could not limit payments for education-related expenses.69 In reaching this decision, the district court highlighted the ambiguity of the concept of amateurism in intercollegiate sports.70 After multiple appeals, the Supreme Court

---

62 Id.
63 Id.
65 See id. at 461, 473–76 (listing the status of all existing state NIL legislation and the status of such legislation as of March 7, 2021). The six states that already had NIL statutes at this point were California, Colorado, Florida, Michigan, Nebraska, and New Jersey. See id. at 475–76.
67 See id. at 2152; see also id. at 2154 (defining monopsony as a monopoly on the buyer’s side of the market).
68 See id. at 2152.
69 See id. at 2153.
70 See id. at 2152; cf. Oral Argument at 00:15, Nat’l Collegiate Athletic Ass’n v. Alston, 141 S. Ct. 2141 (2021) (No. 20-512), http://www.oyez.org/cases/2020/20-512 (“[T]he distinct character of college sports has been that it’s played by students who are amateurs which is to say they are not paid for their play.”).
of the United States granted certiorari and affirmed the district court’s judgment that the NCAA’s limits on education-related benefits violated antitrust law.\footnote{See Alston, 141 S. Ct. at 2154, 2166.} The Supreme Court entered a unanimous decision against the NCAA and thereby allowed student-athletes to receive education-related benefits.\footnote{See id. at 2166.}

In reaching this decision, the Supreme Court clarified that the NCAA’s argument about maintaining amateurism “as a part of serving [the] societally important non-commercial objective’ of ‘higher education’” was a request for “judicially ordained immunity” from restraints of trade.\footnote{See id. at 2158–59 (alteration in original) (quoting Brief for Petitioner at 3, Nat’l Collegiate Athletic Ass’n v. Alston, 141 S. Ct. 2141 (2021) (No. 20-512)); see also id. at 2164 (explaining the NCAA’s concern that the ability to receive paid post-eligibility internships will allow sneaker companies or auto dealerships to offer internships with “extravagant salaries as a ‘thinly disguised vehicle’ for paying professional-level salaries”) (quoting Brief for Petitioner at 37–38, Nat’l Collegiate Athletic Ass’n v. Alston, 141 S. Ct. 2141 (2021) (No. 20-512)).} The Court highlighted that college athletics is a “massive business” generating profit for those who run the enterprise.\footnote{Id. at 2150 (“At the center of this thicket of associations and rules sits a massive business.”).} The Court explained that the NCAA cannot be exempt from restrictions merely because of the overlap between education, sports, and money.\footnote{Id. at 2159.} The Court also foreshadowed possible future legislation by mentioning that the NCAA could argue for exemption from antitrust laws, but that argument should be addressed to Congress.\footnote{Id. at 2160.}

Justice Kavanaugh’s concurring opinion indicated that although the majority’s opinion is limited to education-related benefits, the rest of NCAA’s limits on compensation are subject to the same scrutiny.\footnote{See id. at 2167 (Kavanaugh, J., concurring); see also Advisory Opinions, Supreme Court Rules Against NCAA, SPOTIFY, at 15:18 (June 21, 2021), http://open.spotify.com/episode/0ngChUL54Kpwe9eMzIlZ4h4Jt5i=ce0e6ce66e2433e (mentioning that Justice Kavanaugh is a basketball coach to help explain his strong opinion in favor of college athletes). In fact, in June of 2021, after the Alston decision was released, in House v. NCAA, a California district court refused to dismiss an action opposing the NCAA’s rules restricting student-athletes’ NIL income. See Grant House v.}
Chapman Law Review

C. Further Changes to Paying Student-Athletes since NCAA v. Alston

Since the NCAA v. Alston decision was released, the NCAA has loosened its restrictions on student-athletes’ income, more states have introduced and enacted legislation protecting student-athletes’ ability to profit off their NIL, numerous federal bills have been introduced to create federal legislation for student-athletes’ NIL income, and courts have maintained their reluctance to allow the NCAA to continue to restrict student-athletes’ income.

1. NCAA Interim NIL Policy

Following NCAA v. Alston and the warning set forth in Justice Kavanaugh’s concurring opinion, the NCAA publicized an interim policy allowing NIL payments to student-athletes and permitting them to sign with agents without tarnishing their amateur status.80 This policy allows for college-athletes to engage in NIL activities pursuant to applicable state law.81 Student-athletes...
within states without such laws can engage in NIL activities as long as they do not violate the NCAA’s rules. The announcement claimed that this policy preserved the principle that “college sports are not pay-for-play.” Further, the NCAA clarified its intent to actively work with Congress to create legislation to “support student-athletes.” The NCAA subsequently clarified that the interim policy does not allow for pay-for-play or improper inducement of student-athletes.

2. NIL Legislation

Additionally, since NCAA v. Alston was handed down, more legislation has developed to protect student-athletes’ rights to profit off their NIL. The following chart shows the status of current NIL legislation throughout the United States. The chart includes four categories of states: (1) states with legislation that is signed into law and effective, (2) states with legislation...
that is signed into law, but not yet effective, 89 (3) states with legislation that has been introduced, but not signed into law, 90 and (4) states with no NIL legislation. 91

89 See MD. CODE ANN., EDUC. § 15-131 (LexisNexis 2022) (effective July 1, 2023); MONT. CODE ANN. § 20-1-232 (2021) (effective June 1, 2023); N.J. REV. STAT. § 18A:3B-87 (2022) (effective as of Sept. 14, 2020, but not applicable until the fifth academic year thereafter).


91 The states that never introduced any NIL legislation are not included in the list of statutes introduced and signed into law. See Tracker: Name, Image and Likeness by State, supra note 86. It is noteworthy that Alabama had an NIL legislation bill that was repealed. See H.B. 404, 2021 Leg., 2021 Reg. Sess. (Ala. 2021) (repealed 2022); see also John H. Glenn, Alabama House Passes Repeal Bill for “Restrictive” NIL Law for Student-Athletes, ALA. POL. REP. (Jan. 19, 2022, 2:32 PM), http://www.alreporter.com/2022/01/19/alabama-house-passes-repeal-bill-for-restrictive-nil-law-for-student-athletes/ [http://perma.cc/MP4K-UW4B] (explaining that the bill was repealed to give student-athletes more freedom to profit off their NIL rights because the bill was more restrictive than the NIL standards). Additionally, Kansas and West Virginia had bills that died before they were signed into law. H.R. 2264, 2021 Leg., Reg. Sess. (Kan. 2021) (introduced and approved by the House, but died on Senate General Orders); H.R. 2583, 85th Leg., Reg. Sess. (W. Va. 2021) (introduced, but died in committee).
<table>
<thead>
<tr>
<th>Signed into law and effective</th>
<th>Signed into law, but not yet effective</th>
<th>Introduced, but not signed into law</th>
<th>No NIL legislation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arizona</td>
<td>Maryland</td>
<td>Hawaii</td>
<td>Alabama</td>
</tr>
<tr>
<td>Arkansas</td>
<td>Montana</td>
<td>Iowa</td>
<td>Alaska</td>
</tr>
<tr>
<td>California</td>
<td>New Jersey</td>
<td>Massachusetts</td>
<td>Delaware</td>
</tr>
<tr>
<td>Colorado</td>
<td>Minnesota</td>
<td>Idaho</td>
<td></td>
</tr>
<tr>
<td>Connecticut</td>
<td>New Hampshire</td>
<td>Indiana</td>
<td></td>
</tr>
<tr>
<td>Florida</td>
<td>New York</td>
<td>Kansas</td>
<td></td>
</tr>
<tr>
<td>Georgia</td>
<td>Rhode Island</td>
<td>North Dakota</td>
<td></td>
</tr>
<tr>
<td>Illinois</td>
<td>Vermont</td>
<td>South Dakota</td>
<td></td>
</tr>
<tr>
<td>Kentucky</td>
<td>Washington</td>
<td>Utah</td>
<td></td>
</tr>
<tr>
<td>Louisiana</td>
<td></td>
<td>West Virginia</td>
<td></td>
</tr>
<tr>
<td>Maine</td>
<td></td>
<td>Wisconsin</td>
<td></td>
</tr>
<tr>
<td>Michigan</td>
<td></td>
<td>Wyoming</td>
<td></td>
</tr>
<tr>
<td>Mississippi</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missouri</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nebraska</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nevada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Mexico</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Carolina</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ohio</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oklahoma</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oregon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pennsylvania</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Carolina</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tennessee</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Texas</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Virginia</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A majority of states have NIL legislation signed into law. Oregon enacted additional legislation requiring makers of team jerseys, video games, and trading cards to pay royalties to student-athletes.\(^92\) Taking a different approach, Alabama repealed its NIL statute in an effort to give student-athletes more freedom to earn NIL income; the state’s statute was more restrictive than the NIL legislation.

\(^92\) See S. 1505, 81st Leg., Reg. Sess. (Or. 2022) (effective July 1, 2022).
State NIL legislation is rapidly changing, therefore this table will likely evolve as time passes.94

Additionally, many bills have been proposed to create federal NIL legislation.95 These bills are all substantively similar, with a few distinct differences. It has been ten months since NCAA v. Alston and since then, numerous bills have developed with varying details; it is likely that there will continue to be movement towards federal NIL legislation in the coming months.

II. FEDERAL TAXATION OF STUDENT-ATHLETES

Although taxation of student-athletes is evolving rapidly, in some ways, taxation of professional athletes can be used as a good

---

93 See H.R. 404, 2021 Leg., 2021 Reg. Sess. ( Ala. 2021) (repealed 2022); see also John H. Glenn, Alabama House Passes Repeal Bill for “Restrictive” NIL Law for Student-Athletes, ALA. POL. REP. (Jan. 19, 2022, 2:32 PM), http://www.alreporter.com/2022/01/19/alabama-house-passes-repeal-bill-for-restrictive-nil-law-for-student-athletes/ [http://perma.cc/MP4K-UW4B] (explaining that the bill was repealed to give student-athletes more freedom to profit off their NIL rights because the bill was more restrictive than the NIL standards).

94 Between March of 2021 and March of 2023, the number of state legislatures that passed NIL statutes increased from six to twenty-nine. Compare Changing the Face of College Sports, supra note 64, at 473–76, with infra Part I.C.2.

95 See Tracker: Name, Image, and Likeness by State, supra note 86 (listing federal NIL legislation that has been introduced). The first bill—introduced by Representative Mark Walker in 2019—would alter the definition of a tax-exempt sports organization to prevent the organizations from restricting student-athletes’ NIL rights. See H.R. 1804, 116th Cong. (2019). On June 18, 2020, Senator Marco Rubio introduced a bill which would require an intercollegiate athletic association to establish a policy allowing student-athletes to profit from their NIL and to hire agents to represent them. See S. 4004, 116th Cong. (2020). This bill seems remarkably similar to the NCAA’s interim policy, which was established twelve days after this bill was introduced. Compare S. 4004, 116th Cong. (2020), with NCAA Interim Policy, supra note 80. After the publication of the NCAA v. Alston decision, many more federal bills were released. By September of 2020, Representative Anthony Gonzalez had introduced a bill that would prohibit higher education institutions and athletic organizations from restricting student-athletes’ ability to enter into an endorsement contract or agency contract and would assign enforcement of the policy to the Federal Trade Commission. See H.R. 8382, 116th Cong. (2020). In December of 2020, three more bills were introduced. See S. 5003, 116th Cong. (2020); H.R. 9033, 116th Cong. (2020); S. 5062, 116th Cong. (2020). Senator Roger Wicker introduced a bill that would add a more uniform framework to the NIL compensation structure. See S. 5003, 116th Cong. (2020). Representative Janice D. Schakowsky and Senator Cory A. Booker created a similar comprehensive “College Athletes Bill of Rights” which outlines the student-athletes’ rights to use their NIL, to receive compensation from universities for education-related expenses, to be represented by an agent, to transfer between universities, and to lay claim to other rights and policies that are currently governed by the NCAA. See S. 5003, 116th Cong. (2020); S. 5062, 116th Cong. (2020). At the start of 2021, two more bills were introduced. See S. 414, 117th Cong. (2021); S. 238, 117th Cong. (2021). Senator Jerry Moran introduced a bill that aligned more closely with the NCAA’s goal to limit student-athletes from receiving compensation from NIL deals with unaffiliated third parties. See S. 414, 117th Cong. (2021). Senator Christopher Murphy also introduced a bill that prevents higher education from limiting compensation offered to student-athletes under NIL contracts; regulates athlete representation; authorizes grants for analyzing student-athlete NIL monetization; provides that the Federal Trade Commission is to regulate the policy; and preempts more restrictive state NIL laws. See S. 238, 117th Cong. (2021).
model. Professional athletes, who used to receive preferential tax treatment, have faced heightened scrutiny on their taxes because of their public image. Similarly, with the newly allowed NIL payments, student-athletes are becoming increasingly popular and known throughout the country and the world. It is likely that with the new fame from NIL payments, student-athletes will face heightened scrutiny similar to that faced by professional athletes. A differentiating factor between the tax treatment of professional athletes and student-athletes is that professional athletes are employees of the team for which they play, whereas student-athletes are not. However, even if professional athletes are employees of their athletic organization, this does not mean they are employees for purposes of their NIL deals.

Similar to how professional athletes became bigger tax targets as they were paid more, it is likely that because of the newly allowed NIL payments, student-athletes will become more highly scrutinized as taxpayers. However, unlike professional athletes, student-athletes are young and inexperienced in paying taxes or handling finances. Student-athletes are as young as seventeen; they cannot even legally vote, purchase alcohol, or purchase cigarettes, yet they are expected to keep track of, calculate, and pay complicated taxes. Professional athletes,

---


97 See Kathryn Kisska-Schulze & Adam Epstein, “Show Me the Money!”—Analyzing the Potential State Tax Implications of Paying Student-Athletes, 14 VA. SPORTS & ENT. L.J. 13, 29 (2014) [hereinafter Show Me the Money] (discussing the heightened scrutiny of taxation of professional athletes); see also John DiMascio, The “Jock Tax”: Fair Play or Unsportsmanlike Conduct, 68 UNIV. PITTS. L. REV. 953, 957 (2007) (recognizing that it is more difficult for professional athletes to avoid their taxes because of their public image and heightened popularity); see also John T. Holden & Kathryn Kisska-Schulze, Taxing Sports, 71 AM. UNIV. L. REV. 845, 856 (2022) [hereinafter Taxing Sports] (mentioning that the NFL enjoyed tax-exempt status as being not-for-profit between 1942 and 2015).


99 See id. at 4:35 (highlighting that professional athletes have taxes withheld because they are employees, whereas college athletes are not employees and as such, will need to be more responsible than professional athletes); see also infra Part II.B.1.

who have more experience with NIL income and the related tax
effects, likely also have assistance with the calculation, filing,
and proper payment of their taxes. Student-athletes must be
assisted to ensure they avoid penalties for dropping the ball on
the compliance of their tax returns.

Because, historically, student-athletes were unable to receive
most forms of income, student-athletes generally did not need to
file taxes (unless they were also working while in college). With
the newly allowed NIL income, as soon as a student-athlete earns
$12,550, they will need to file a federal tax return.101

A. Evolution of Taxation of Student-Athletes and the Decrease
of Preferential Treatment

Since 1954, college and university students have received
many forms of tax benefits.102 Among others, these include a tax
exclusion for reductions of tuition,103 a tax deduction for interest
paid on student loans,104 and a tax exemption for parents of college
students between the ages of nineteen and twenty-three.105
Additionally, the Internal Revenue Code currently excludes any
amount received as a “qualified scholarship” from gross income for
tax purposes.106 A “qualified scholarship” includes any money
received as a scholarship used for tuition, required fees, books,
supplies, and equipment required for courses.107

In recent decades, college athletics has received preferential
treatment for federal tax purposes.108 Since 1976, college athletics
organizations and the NCAA have enjoyed tax exempt status
because of their goal “to foster national or international amateur
sports competition.”109 For the same reason, any donation to either

101 See U.S. DEP’T OF TREAS., IRS, PUB. NO. 501, DEPENDENTS, STANDARD DEDUCTION,
AND FILING INFORMATION: FOR USE IN PREPARING 2021 RETURNS, at 2–4 (2022) (applying
even if the college-athlete is still a minor); see also Probasco, supra note 100.
102 See MARGOT L. CRANDALL-HOLICK, CONG. RSCH. SERV., R41967, HIGHER
EDUCATION TAX BENEFITS: BRIEF OVERVIEW AND BUDGETARY EFFECTS 2 (2021),
103 See I.R.C. § 117.
104 See I.R.C. § 221.
105 See I.R.C. §§ 151(c), 152 (c)(3)(ii).
106 See I.R.C. § 117(a).
107 See I.R.C. § 117(b).
108 See Kathryn Kisska-Schulze, This is Our House!—The Tax Man Comes to College
Sports, 29 MARQ. SPORTS L. REV. 347, 348 (2019) [hereinafter The Tax Man Comes to College
Sports] (“Prior to 2018, college athletics had historically enjoyed favorable federal tax
treatment due to the tax-exempt status of universities, athletic departments, and the National
Collegiate Athletic Association (NCAA).”); see also Kathryn Kisska-Schulze & Adam Epstein,
The Claim Game: Analyzing the Tax Implications of Student-Athlete Insurance Policy Payouts,
25 JEFFREY S. MOORAD SPORTS L.J. 231, 250 (2018) (“[T]he college sports industry in general
has enjoyed strikingly favorable tax positions over the past decades due to the tax-exempt
status of colleges, athletic departments, and the NCAA itself.”).
109 See I.R.C. § 501(c)(3).
college athletic departments or the NCAA is considered a charitable donation and is deductible from the donor’s taxable income (within prescribed limits). Another benefit college athletics enjoyed before 2017 was that eighty percent of amounts paid for the rights to purchase seats at college athletics events were deductible as charitable contributions. In 2017, the Tax Cuts and Jobs Act (“TCJA”) was rapidly enacted, which altered the historic preferential treatment of college athletics. Mainly, the TCJA repealed the deduction for amounts paid in exchange for college athletic event seating rights.

Additionally, the I.R.C. has traditionally qualified student-athletes’ scholarships under the scholarship tax exclusion. The IRS has maintained the exclusion for college athletic scholarships because of the absence of any quid pro quo relationship. Courts have held that scholarship and grant funds qualify for the exclusion so long as those funds are not received in exchange for services rendered. With the new ability for student-athletes to...
receive education-related benefits from their college or university, will this pay be included in the exclusion for qualified scholarships? According to the Supreme Court's interpretation of the statute for student athletic scholarships, they would still fall under the broad umbrella of the exception because the education-related benefits are not in exchange for services rendered.  

With the new ability to earn compensation from NIL deals, the applicability of the scholarship tax exclusion for athletic scholarships has been challenged. In contrast, the NCAA clarified that its NIL interim policy "preserve[d] their commitment to avoid pay-for-play." Senator Richard Burr introduced a bill that would deny the tax exclusion for athletic scholarships for student-athletes that make more than $20,000 of profit during the taxable year from their NIL. NIL income is unrelated to a student-athlete's athletic scholarship — the NIL income is from a third party, completely unrelated to the university and does not interfere with the avoidance of the pay-for-play model. So, the NIL income should not affect the ability to exclude the scholarships. This challenge to the traditional preferential tax treatment for student-athletes is likely a benchmark for the beginning of a trend toward more heightened scrutiny in taxing student athletics.

B. Complications with Independent Contractor Status

Independent contractors and employees have different rights and obligations, so it is important to determine whether a student-athlete is an independent contractor or an employee of the companies paying them pursuant to a NIL contract. Employers have an incentive to consider student-athletes as independent contractors to avoid vicarious liability, avoid paying them benefits, avoid issues regarding workers' compensation, and avoid their share of employment taxes. Although independent contractors are unable to utilize employee benefits, workers' compensation, or lower employment taxes, they may deduct ordinary and necessary

---

his services rendered). But see Bingler v. Johnson, 394 U.S. 741, 757–58 (1969) (holding that a Ph.D. student's stipend did not qualify for the Section 117 exclusion because it was compensation for services, rather than a scholarship).

117 See Bingler, 394 U.S. at 757–58.

118 See Taxing Sports, supra note 97, at 906.

119 See NCAA Interim Policy, supra note 80.

120 See Taxing Sports, supra note 97, at 869–70.

121 See id. at 907; NCAA Interim Policy, supra note 80.

122 See Taxing Sports, supra note 97, at 857 (acknowledging that some of the "wide berth of favorable tax treatment . . . is now shifting").

123 See, e.g., RESTATEMENT (SECOND) OF AGENCY § 219 (AM. L. INST. 1958) (explaining that employers can be vicariously liable for the torts of their employees); see also I.R.C. §3402(a) (requiring employers to withhold employment taxes from the pay of employees).
business expenses. Student-athletes making meager amounts of money from their NIL will benefit more from using the standard deduction because they will be able to minimize their taxes, without the administrative burden of keeping track of their ordinary and necessary business expenses.

1. Student-Athletes Will Generally Be Independent Contractors

An independent contractor provides a service for the hiring party but is neither controlled nor subject to the hiring party’s right to control the physical conduct of performing the task. In that case, the hiring party has the right to control the end result of the work, but not how the work will be done. In contrast, where there is an employer-employee relationship, the employee is subject to the employer’s right to control the physical conduct of performing the service. Whether a worker is an independent contractor or an employee is a fact-intensive determination which must be decided on a case-by-case basis. Among other factors, the IRS considers the degree of control, the investment in facilities/equipment, the length of the relationship, and the parties’ subjective belief of the worker’s classification. It is likely that companies paying student-athletes for their NIL will classify them as independent contractors, but their subjective belief about the nature of the relationship is not controlling; it is merely one factor in the fact-intensive determination.
determination. It is likely that most students with NIL deals will be considered independent contractors.

2. Ordinary and Necessary Business Expense Deductions

Certain student-athletes may be able to deduct certain expenses as ordinary and necessary business expenses. Taxpayers may deduct "all the ordinary and necessary expenses paid or incurred during the taxable year in carrying on any trade or business." To reap the benefits of this deduction, the expense must be ordinary, necessary, and incurred in carrying on the taxpayer's trade or business.

To qualify as carrying on a trade or business, the taxpayer must be involved in the activity with continuity and regularity and be engaged in the activity for the primary purpose of income or profit. In contrast, a mere hobby or sporadic activity will not qualify. A professional athlete engages in a trade or business when they use NIL in exchange for endorsement income. While the ultimate determination depends on the facts of each case, the

---

131 See id.; McCann & Raiola, supra note 124.
132 See Jeremy Crabtree, College Players Who Made NIL Money Have New Homework—Paying Taxes, ON3 NIL (Apr. 18, 2022), http://www.on3.com/nillnews/college-players-who-made-nil-money-have-new-homework-paying-taxes/ [http://perma.cc/7KFV-KJV3]. In NIL deals, it is unlikely that the employer will have the right to control how the student-athlete finishes the work. See generally See Common-Law Employee, supra note 128 (defining employee and using right to control as the determinative element); Independent Contractor Defined, supra note 126 (defining independent contractor). Additionally, (1) the employer probably does not have control over the details of student-athlete work; (2) student-athletes will use their own phones/equipment for their NIL deals and will provide their own place of work; (3) the opportunity for profit or loss will likely range from deal to deal, but student-athletes will likely get more income if, for example, their promotion of products attracts more costumers; (4) the company paying for student-athletes' NIL likely has a right to end the partnership; (5) advertising is usually a part of businesses, but not usually an integral part of the business (this factor is a close call); (6) some NIL deals are a one-time thing, while others span multiple years, depending on the particular deal; (7) the parties will likely believe the student-athletes are independent contractors; and (8) NIL deals will not likely include any employee benefits. See generally Keller, 103 T.C.M. (CCH) at 1298 (listing these factors for how to determine whether a worker is an employee or independent contractor). Other than the fact that companies likely have the right to terminate their relationship with student-athletes, the factors lean toward student-athletes being classified as independent contractors. However, it seems possible for a student-athlete to be considered an employee where the athlete has a long-term NIL deal, the athlete is paid for time spent, and the hiring party has control over the details and the physical aspect of the work.
133 I.R.C. § 162(a). Before 2018, Section 212 of the I.R.C. provided a deduction for “all ordinary and necessary expenses paid or incurred during the taxable year for the production or collection of income,” but the deduction is not allowed to be taken for the taxable years starting on December 31, 2017, until January 1, 2026. See I.R.C. § 212; see also I.R.C. § 67(g).
134 See I.R.C. § 162(a); see also Bobby L. Dexter, Federal Income Taxation in Focus 341 (Rachel E. Barkow et al. eds., 2d ed. 2022).
136 See id.
137 See Changing the Face of College Sports, supra note 64, at 488–89.
IRS will likely view such activity as a trade or business.¹³⁸ NIL deals with longer terms and more continuous work are more likely to qualify as a trade or business.¹³⁹ In addition, using a business entity could further the argument that a student-athlete’s NIL income is a trade or business because the development of the business entity creates structure and shows that the income is more than a mere hobby.¹⁴⁰

Conversely, student-athletes’ focus while in college is on their education and athletics. Because student-athletes’ income is not classified as pay-for-play, the income is not in exchange for their work as a student-athlete, it could be considered a mere hobby.¹⁴¹ Especially for a student athlete with a single, short-term NIL deal (possibly even a NIL deal for a single social media post or a single advertisement), the argument that the pursuit of NIL income is a mere hobby is stronger.¹⁴² However, the student-athletes’ NIL only has value because of their performance on the field.

Treasury regulations provide some guidance for endorsement contracts.¹⁴³ The treasury regulations clarify that a well-known chef and restaurant owner who receives endorsement income due to their skill and reputation is in the trade or business of being a chef, owning a restaurant, and earning endorsement fees.¹⁴⁴ Similarly, a well-known actor who enters into a partnership with a shoe company where the actor contributes their NIL in exchange for fifty percent of the income in the partnership is in the trade or business of receiving partnership interest in exchange for their NIL.¹⁴⁵ Following the guidance set forth in these treasury regulations, student-athletes’ pursuit of NIL income will also be

---

¹³⁸ See id.; see generally Groetzinger, 480 U.S. at 35–36 (setting forth the Groetzinger standard for distinguishing a trade or business from a mere hobby).
¹³⁹ See generally Groetzinger, 480 U.S. at 35–36 (setting forth the Groetzinger standard for distinguishing a trade or business from a mere hobby).
¹⁴¹ See generally NCAA Interim Policy, supra note 80 (clarifying that NIL income is not pay-for-play). Prior to the Tax Cuts and Jobs Act, I.R.C. Section 212 allowed individuals to take a deduction for ordinary and necessary expenses incurred in the production of income from an activity that does not constitute a trade or business, however, this Section is disallowed for the taxable years from December 31, 2017, to January 1, 2026. See I.R.C. §§ 67(g), 212.
¹⁴² See generally Groetzinger, 480 U.S. at 35–36 (setting forth the Groetzinger standard).
¹⁴³ See Treas. Reg. § 1.199A-5(b)(3)(xv)–(xvi) (2019). This treasury regulation relates to another deduction. However, a trade or business within the meaning of IRC Section 199A has the same meaning as “a trade or business that is a trade or business under Section 162 . . . other than the trade or business of performing services as an employee.” See Treas. Reg. § 1.199A-1(b)(14) (2019). Because student-athletes are not employees, for the purposes of this article, the use of “trade or business” within I.R.C. Section 199A and I.R.C. Section 162 are synonymous.
considered a trade or business. Ultimately, the qualification of the student-athletes’ pursuit of NIL income as a trade or business will be a case-by-case determination based on the terms of the NIL deals, but it is likely that most student-athletes’ pursuit of NIL income will be considered a trade or business.146

Student-athletes must also determine which expenses may be deducted as ordinary and necessary business expenses. To be “ordinary,” incurring and paying such expense must be the “common and accepted means” of addressing a given business situation in the taxpayer’s community.147 To be “necessary,” the expense must be “appropriate and helpful.”148 Before a student-athlete can take a deduction, they must determine if the expense, according to these definitions, is ordinary and necessary. Among the items allowed to be deducted are supplies, management expenses, advertising, and traveling expenses while away from home solely in the pursuit of a trade or business.149 For student-athletes generating NIL income, this could include cameras for vlogging or recording unboxing videos of sponsored products and traveling to sports camps if the camps are in a different location than their tax home. If they hire an agent to help obtain NIL deals and manage the business, then payments to the agent (and many other ordinary and necessary expenses incurred in producing their NIL income) likely qualify. However, even if an expense would qualify as ordinary and necessary, the deduction would not be allowed for any expenses that are already covered by the student-athlete’s college or university (such as traveling to compete in a sport on behalf of the school).150

For traveling expenses, student-athletes will need to determine their tax home because travel expenses are only deductible “while away from home.”151 Generally, a taxpayer’s home for the purposes of Section 162 of the I.R.C. is their principal place of business.152 This becomes more complicated when a taxpayer lives at multiple residences.153 Circuits are split

146 See Changing the Face of College Sports, supra note 6464, at 488–89 (concluding that NIL income is derived from a trade or business). See generally Groetzinger, 480 U.S. at 35–36 (setting forth the Groetzinger standard).
147 See Welch v. Helvering, 290 U.S. 111, 114 (clarifying that an expense was ordinary because it was “the common and accepted means” to respond to the situation at hand); see also Gilliam v. Comm’r, 51 T.C.M. (CCH) 515 (1986) (describing ordinary as “normal, usual, or customary” and explaining that “it must be of common or frequent occurrence in the type of business involved”).
148 See Welch, 290 U.S. at 113 (finding that an expense was necessary because it was “appropriate and helpful”).
150 See Changing the Face of College Sports, supra note 64, at 490.
151 See I.R.C. § 162(a)(2).
152 See Markey v. Comm’r, 490 F.2d 1249 (6th Cir. 1974).
153 See Changing the Face of College Sports, supra note 64, at 491.
as to whether a taxpayer’s home is tied to their principal place of business or their actual residence. 154 When professional athletes play a team sport, their tax home is usually where their team is located. 155 Professional athletes who play an individual sport, such as swimming or gymnastics, have more flexibility in deciding their tax home. 156 However, student-athletes are more similar to professional athletes that play a team sport because, even if a student-athlete plays an individual sport, they are required to attend school, practice, and live a majority of the year where their school is located.

This deduction is helpful but is also extremely burdensome. To enjoy the benefits of this deduction, student-athletes would need to determine their tax home, track all their expenses related to their production of NIL income, and determine whether each expense is ordinary and necessary. 157

3. Employment Taxes

Generally, employers are required to withhold employment taxes from employees’ wages and pay those taxes to the IRS. 158 Wages include “all remuneration for employment.” 159 Employment taxes include old age, survivors, and disability insurance (“OASDI”) taxes and hospital insurance (“HI”) taxes, for a total of 7.65% of wages being withheld (plus another 0.9% for HI taxes for wages after a certain threshold). 160 There is a cap on OASDI employment taxes called the contribution and benefit base, which is commonly referred to as the FICA Wage Base. 161 In practice, after the employment taxes are withheld on the wage base, the remaining wages are not subject to OASDI withholding.
but they are still subject to HI tax withholding.\textsuperscript{162} The current wage base is $147,000.\textsuperscript{163} Additionally, employers are responsible for another 7.65\% of their employee’s wages for their half of employment taxes.\textsuperscript{164}

As discussed above, student-athletes are not employees.\textsuperscript{165} Employers are only obligated or authorized to withhold taxes from employees,\textsuperscript{166} therefore none of a student-athlete’s income will be subject to withholding by an employer. However, student-athletes will still be responsible for employment taxes if they make $400 or more during the taxable year.\textsuperscript{167} Those who are self-employed are responsible for double the amount of employment taxes on their self-employment income.\textsuperscript{168} As self-employed individuals, they are required to pay 12.4\% of self-employment income for OASDI taxes, 2.9\% of self-employment income for HI taxes, and an additional 0.9\% of self-employment income for HI taxes after a certain income threshold.\textsuperscript{169} An individual’s self-employment income includes “the gross income derived by an individual from any trade or business carried on by such individual, less the deductions . . . attributable to such trade or business.”\textsuperscript{170} Unlike employees, self-employed individuals calculate employment taxes after calculating any deduction for their business.\textsuperscript{171} Therefore, self-employed individuals pay employment taxes on a slightly lesser portion of their total income than employees. Like employees, self-employed individuals benefit from the same wage base limits on the OASDI tax.\textsuperscript{172} Additionally, because student-athletes’ employment taxes are not withheld before dispersing payments to them, the student-athlete will need to pay these taxes, in addition

\begin{footnotesize}
\begin{enumerate}
\item See id.
\item See I.R.C. § 3111(a) (requiring employers to pay 6.2\% of employees’ wages for OASDI taxes); I.R.C. § 3111(b) (requiring employers to pay 1.45\% of employees’ wages for HI taxes). Employers are not subject to the same 0.9\% increase as employees after a certain threshold income. \textit{Compare} I.R.C. § 3111(a)–(b), \textit{with} I.R.C. § 3101(a)–(b).
\item See \textit{supra} Part II.B.1.
\item See I.R.C. § 3402(a).
\item See I.R.C. § 6017.
\item \textit{Compare} I.R.C. § 1401(a)–(b), with I.R.C. § 3111(a)–(b), and I.R.C. § 3101(a)–(b).
\item See I.R.C. § 1401(a)–(b).
\item \textit{See} I.R.C. § 1402(a). Here, trade or business has the same meaning as trade or business in I.R.C. Section 162 for ordinary and necessary business expenses. \textit{See} I.R.C. § 1402(b). For an explanation on why student-athletes’ use of their NIL for compensation is a trade or business, \textit{see supra} Part II.B.2. (determining that student-athletes’ use of their NIL for compensation likely satisfies the “trade or business” requirement for ordinary and necessary business expenses).
\item \textit{See} I.R.C. § 1402(a).
\item \textit{See} I.R.C. § 1402(b)(1); \textit{see also} Contribution and Benefit Base, SOC. SEC., \url{http://www.ssa.gov/oact/cola/cbb.html} [http://perma.cc/MSUL-72FU] (last visited Apr. 23, 2022).
\end{enumerate}
\end{footnotesize}
to their normal taxes, at the same time. Rather than spreading the burden of the employment taxes over payments, they will have to pay large amounts in lumped sums, which will be added to any other taxes the student-athlete may already owe. Luckily, to ease the burden, the I.R.C. allows self-employed individuals to deduct one-half of their employment taxes. Although it is helpful, this deduction does not completely eradicate the extra burden on self-employed individuals because a deduction merely decreases the amount of taxable income, rather than reducing the amount of tax owed. Further, self-employment taxes are only imposed on 92.35% of an individual’s net income.

4. Quarterly Estimated Tax Payment

As self-employed individuals, student-athletes will be required to make quarterly estimated payments. This requires student-athletes to not only become familiar with their finances and taxes for tax day, but also to calculate their taxes four times a year to get an accurate estimate to make their quarterly tax payment.

C. Using the Limited Liability Company Business Form to Minimize Taxes

Student-athletes could consider using business organizations to avoid personal liability if they get sued for torts or other claims. Avoiding personal liability will be especially beneficial for student-athletes organizing clinics and camps. In deciding what kind of business entity to form, it is important to consider taxes. The options for business organization types include: sole proprietorship, partnership, limited liability partnership, limited partnership, limited liability limited partnership, limited liability

175 See I.R.C. § 1402(a)(12) (allowing self-employment income to be reduced by 7.65%); see also Changing the Face of College Sports, supra note 64, at 494.
177 See Wittry, supra note 177. This limited liability “does not apply to personal tortious acts or personal guarantees of corporate obligations.” J. DENNIS HYNES & MARK J. LOEWENSTEIN, AGENCY, PARTNERSHIP, AND THE LLC: THE LAW OF UNINCORPORATED BUSINESS ENTERPRISES 4 n.3 (10th ed. 2019).
179 See HYNES & LOEWENSTEIN, supra note 177, at n.4.
company, corporation, and S corporation.\textsuperscript{180} Out of these options, a limited liability company is the best choice for student-athletes.

A sole proprietorship is created when an individual owns an unincorporated business by themselves.\textsuperscript{181} This is a default business organization;\textsuperscript{182} if the student-athletes profit from their NIL by themselves, they would be viewed as sole proprietors. However, a sole proprietorship does not provide the limited liability shield from which a student-athlete would benefit.\textsuperscript{183} Similarly, regardless of intent, a partnership is formed when an association of at least two people (human beings or business entities) carry on, as co-owners, a business for profit.\textsuperscript{184} Therefore, if the student-athlete works with another person (perhaps a family member or the school the student-athlete attends) to create profit on the student-athlete’s NIL, they could default into a partnership. Like a sole proprietorship, partnerships do not have the benefit of limited liability for the partners.\textsuperscript{185}

It would be advantageous for the student-athlete to avoid the possibility of unintentionally forming a partnership and to limit liability by creating a different type of business organization. One option is a Limited Partnership, but a Limited Partnership would also have drawbacks. A Limited Partnership has at least two partners: a general partner and a limited partner.\textsuperscript{186} The general partner has complete control of the Limited Partnership, but can also be held personally liable for the debts and obligations of the Limited Partnership.\textsuperscript{187} The limited partner enjoys the protections of limited liability, but has no control over the Limited Partnership.\textsuperscript{188} If the limited partner exhibits control over the partnership, they could become personally liable for the debts and obligations of the Limited Partnership as a general partner.\textsuperscript{189}

\textsuperscript{180} See id. at 4–10.
\textsuperscript{181} See id. at 6.
\textsuperscript{182} See id.
\textsuperscript{183} See id.
\textsuperscript{184} See, e.g., UNIF. P'SHIP. ACT § 6 (NAT'L CONF. OF COMM'R'S ON UNIF. STATE L. 1914); see also REVISED UNIF. P'SHIP. ACT § 202(a) (NAT'L CONF. OF COMM'R'S ON UNIF. STATE L. 2013); see also CAL. CORP. CODE § 16101(9) (West 2020).
\textsuperscript{185} See HYNES & LOEWENSTEIN, supra note 177, at 9.
\textsuperscript{186} See id. at 10.
\textsuperscript{187} See id.
\textsuperscript{188} See id.
\textsuperscript{189} See UNIF. LTD. P'SHIP. ACT § 7 (NAT'L CONF. OF COMM'R'S ON UNIF. STATE L. 1916) (holding a limited partner liable as a general partner if the limited partner takes control in the business) (amended 1976); see also REVISED UNIF. LTD. P'SHIP. ACT § 303(a) (NAT'L CONF. OF COMM'R'S ON UNIF. STATE L. 1976) (holding a limited partner liable as a general partner if the limited partner exhibits substantially the same amount of control in the business as a general partner or exhibits control and transacts with a third party who has actual knowledge of the limited partner’s participation in the control) (amended 1985); REVISED UNIF. LTD. P'SHIP. ACT § 303(a) (NAT'L CONF. OF COMM'R'S ON UNIF. STATE L. 1985)
Although there is an opportunity for limited liability in a Limited Partnership, it is not the best choice for a student-athlete. This is because—in order to enjoy the benefits of limited liability—a student-athlete would need someone trustworthy who is willing to assume personal liability for the debts and obligations of the limited partnership as the general partner, and the student-athlete would need to relegate all control in the business to the general partner. If the student-athlete wanted to remain in control, the student-athlete could create a corporation to act as the general partner; however, there simply is no need to go through that entire process if the goal is merely to limit liability.

The Limited Liability Limited Partnership ("LLLP") is similar to the Limited Partnership, except none of the partners are personally liable for the debts and obligations of the business. LLLPs may be a good choice for student-athletes, but they do not exist in all states. California is one of the states that does not have an LLLP statute. The LLLP began to grow throughout the United States, but is now unnecessary because of the rise of the LLC. Additionally, LLLPs require at least two partners; therefore, a student-athlete would be unable to create an LLLP by themselves without creating an extra business entity.

A corporation would allow the student-athlete to avoid personal liability as the owner of the company but requires certain corporate formalities to be followed and requires an extra layer of taxation. To form a corporation, the student-athlete would need to file the articles of incorporation with the respective Secretary of State. If corporate formalities are not used, the student-athlete risks the possibility of being held personally

---

190 See HYNE & LOEWENSTEIN, supra note 177, at 10.
192 See id.
193 See generally HYNE & LOEWENSTEIN, supra note 177, at 10–11 (summarizing the LLLP and the LLC).
194 See generally id. at 10; see also UNIF. LTD. P'SHIP ACT § 102(11) (UNIF. L. COMM’N 2001) (noting that a limited liability limited partnership is a form of limited partnership, which is an “an entity, having one or more general partners and one or more limited partners, which is formed . . . by two or more persons.”).
195 See id. at 4.
196 See MODEL BUS. CORP. ACT § 2.01 (COMM. ON CORP. L. 2002) (amended 2016).
liable through a piercing the corporate veil claim. The student-athlete would need to elect a board of directors, hold annual shareholder meetings, and keep corporate records. Corporations are more ideal for large businesses that require substantial amounts of capital from selling shares on the public market. Because a student-athlete’s corporation would not likely have many shareholders, student-athletes may be able to benefit from close corporation rules. In order to obtain these benefits, the student-athlete would need to elect to become a close corporation. In some states, a close corporation can avoid certain corporate formalities by including in its articles of incorporation that the corporation will be managed by the stockholders of the corporation rather than the board of directors. Even then, corporations are still subject to double taxation. Certain small corporations can avoid double taxation by using the S corporation classification. However, S corporations are subject to many technical restrictions. An S corporation is defined by the IRS as a domestic corporation, with less than 100 shareholders, only one class of stock, and without any corporate partners, that elects for pass through taxation. An additional downfall to the S corporation entity is that not all states have the S corporation classification.

---

197 See Sea-Land Servs., Inc. v. Pepper Source, 941 F.2d 519, 520–21 (7th Cir. 1991) (using “failure to . . . comply with corporate formalities” as a factor to determine whether there is a unity of interest and ownership of the corporation to justify piercing the corporate veil).
198 See HYNES & LOEWENSTEIN, supra note 177, at 4.
199 See id. at 5.
200 Close corporation statutes vary from state to state. Delaware defines close corporations as corporations with less than thirty shareholders, with only one class of stock, that is not sold on the public market. DEL. CODE ANN. tit. 8, § 342(a) (1953). California defines a close corporation as a corporation whose articles of incorporation contain a provision limiting the amount of shareholders to thirty-five or less and a provision stating, “[t]his corporation is a close corporation.” CAL. CORP. CODE § 158(a) (West 2017).
201 See HYNES & LOEWENSTEIN, supra note 177, at 4.
202 For example, in Delaware, if the Certificate of Incorporation (Delaware’s name for the articles of incorporation) provides that the corporation will be managed by the stockholders of the corporation rather than the board of directors, then annual stockholder meetings are not necessary. DEL. CODE ANN. Tit. 8, § 351 (1953).
203 See HYNES & LOEWENSTEIN, supra note 177, at 4.
204 See I.R.C. § 1361(a)(1)–(b)(1) (defining an S corporation as a corporation with only one class of stock and less than one hundred shareholders—who are all individual residents of the United States—that elects to be an S corporation); I.R.C. § 1366 (explaining how items are passed through to be taxed by the shareholders).
205 See HYNES & LOEWENSTEIN, supra note 177, at n.4.
The Limited Liability Company (“LLC”) is a relatively new type of unincorporated business entity. The first LLC statute was created in Wyoming in 1997 and has since become the most popular business entity in the country. LLCs exist in every state. The LLC combines the tax advantages of a partnership with the limited liability advantages of a corporation. However, the benefits of limited liability, like in corporations, have limits. Similar to the corporate form, with LLCs, if the members do not keep the entity separate from personal activities, they risk being subject to personal liability from a piercing the LLC veil claim.

LLCs allow for flexibility in the management of the business and can be created with only one owner. An LLC can be member-managed or manager-managed. Therefore, if the student-athlete wants to be the sole member of the LLC and have complete control of the business, they have the freedom to do so. On the other hand, if the student-athlete wants to hire someone to manage the LLC (perhaps a parent, agent, accountant, or lawyer), or even manage the LLC with the help of

208 See Hynes & Loewenstein, supra note 177, at 797; see also Elf Atochem N. Am., Inc. v. Jaffari, 727 A.2d 286, 287 (Del. 1999) (“The limited liability company (‘LLC’) is a relatively new entity that has emerged in recent years as an attractive vehicle to facilitate business relationships and transactions.”).

209 See Hynes & Loewenstein, supra note 177, at 797; see also id. at 788 (describing LLCs as “the clear choice for new businesses in the United States”); Rodney D. Chrisman, LLCs are the New King of the Hill: An Empirical Study of New LLCs, Corporations, and LPs Formed in the United States Between 2004-2007 and How LLCs Were Taxed for Tax Years 2002-2006, 15 Fordham J. Corp. & Fin. L. 459, 459–60 (2010) (emphasizing that LLCs are “undeniably the most popular form of new business entity in the United States”); id. at 473–75 (comparing amounts of new LLCs in each state with the number of new corporations and LPs in that state in 2007).

210 See Chrisman, supra note 209, at 473–75 (listing all 50 states and Washington D.C. with each state’s statistics on new LLCs in 2007).

211 See Hynes & Loewenstein, supra note 177, at 797–98 (explaining that owners are not vicariously liable for the obligations of the business and that income of the business passes directly to the owners of the LLC, avoiding being taxed at the entity level); see also Jaffari, 727 A.2d at 287 (“The limited liability company (LLC) is a relatively new entity that has emerged in recent years as an attractive vehicle to facilitate business relationships and transactions.”).

212 See Hynes & Loewenstein, supra note 177, at 817. Piercing the corporate veil also requires some type of injustice which would be prevented by piercing the corporate veil. See Sea-Land Servs., Inc. v. Pepper Source, 941 F.2d 519, 520 (7th Cir. 1991). Additionally, one of the factors to find unity and ownership in piercing the corporate veil claims is avoiding corporate formalities. See Hynes & Loewenstein, supra note 177, at 817. However, to succeed on a piercing the LLC veil claim, corporate formalities are not considered. See id. Thus, piercing the LLC veil would not be identical to a piercing the corporate veil claim. See id.

213 See Hynes & Loewenstein, supra note 177, at 10–11; see also UNIF. LTD. LIAB. CO. ACT § 202(a) (NAT’L CONF. OF COMM’RS ON UNIF. STATE L. 1996) (“One or more persons may organize a limited liability company . . . .” (amended 2013)).


215 See id.
a hired manager, the student-athlete has the freedom to do so.\footnote{216} Any managers hired to manage the LLC are subject to fiduciary duties to ensure that the manager is loyal, discharging their duties with care, and acting in good faith.\footnote{217} If the student-athlete is making small amounts of NIL income and does not have an agent, it may be easy for or a family member to manage the LLC. If a student-athlete is making millions of dollars on their NIL, it may be more beneficial to hire an agent and design the LLC to be managed by the agent so that the student-athlete can focus on their studies and athletics.

The LLC option is more advantageous than a corporation because it does not require burdensome corporate formalities and avoids double-taxation.\footnote{218} Even though S corporations are another way to avoid double-taxation, they still have burdensome restrictions, require a certain level of corporate formality, and do not exist in all states.\footnote{219} Therefore, the LLC is still the more advantageous choice. The LLC is a better option than the partnership because it allows the owners to have limited liability, even if they exert control over the organization.\footnote{220} The only reason the LLC is more burdensome than the partnership is that the LLC is not a default business organization.\footnote{221} However, creating an LLC is not difficult. To create an LLC, the student-athlete (or whoever is creating the LLC on their behalf) would simply need to file the articles of organization with the Secretary of State.\footnote{222} There is generally an affordable fee to file the articles of organization.\footnote{223} An LLC may also be subject to annual fees which are subject to the laws within each state.\footnote{224}

No matter what business entity is right for the student-athlete, there will be pros and cons. The creation of any

\footnote{216 See id.}
\footnote{217 See id. \S 409.}
\footnote{218 See HYNES & LOEWENSTEIN, supra note 177, at 798.}
\footnote{219 See id.}
\footnote{220 See id. at 797–98.}
\footnote{221 See id.}
\footnote{222 See CAL. CORP. CODE \S 17702.01(a) (West 2022). The articles of organization for an LLC only need to include the name of the LLC, the address of the office, the name and address of an agent of the LLC, a statement that the LLC “is to engage in any lawful act,” and in certain circumstances, a statement regarding the management structure of the LLC. See id. \S 17702.01(b).}
\footnote{224 See Limited Liability Company, STATE OF CAL. FRANCHISE TAX BD., http://www.ftb.ca.gov/file/business/types/limited-liability-company/index.html [http://perma.cc/C7Z6-6KXE] (last visited May 14, 2022) (explaining that California LLCs are only subject to annual fees if the LLC makes more than $250,000).}
incorporated business comes with responsibilities. It is hard to imagine an eighteen-year-old right out of high school who is already juggling school and athletics also carrying the added responsibility of a business entity. For many student-athletes earning less NIL income who do not have strong interests in limited liability, it may not even be worth setting up the LLC. However, for most student-athletes, it is likely worth the effort of filing the articles of organization and paying the affordable fee to protect themselves.

III. STATE TAXATION OF STUDENT-ATHLETES

A. In Which State(s) Must Student-Athletes File Taxes?

States have the constitutional power to tax both (1) residents on all sources of income no matter where derived and (2) nonresidents on all income earned from sources within the state. Therefore, student-athletes may have the obligation to file taxes in more than one state.

1. Student-Athletes’ State of Residency

So where are student-athletes legally residing? Each state has different rules for whether an individual is a resident of the state. In California, an individual is a resident when the individual is “in [the] state for other than a temporary or transitory purpose” and when they are “domiciled in [the] State [but are] outside the State for a temporary or transitory purpose.” Additionally, there is a rebuttable presumption that an individual is a resident if they “spend in the aggregate more than nine months of the taxable year within [the] State.” In Texas, an individual is a resident for tax purposes when they live in the state. This includes “[a] person who is temporarily living in the state, and retains a permanent

225 See Wittry, supra note 96 (“[O]nce you go down that LLC, incorporated road . . . there’s [sic] hoops you have to jump through and we can’t ignore those hoops or they’ll come back and bite you.”).
226 California only charges a seventy dollar filing fee for LLCs. See Business Entities Fee Schedule, supra, note 223.
227 See Shaffer v. Carter, 252 U.S. 37, 52, 59 (1920) (holding that states may tax nonresidents on all income derived from a source within the state); see also Show Me the Money, supra note 97, at 46 (“[S]tates have the constitutional power to tax their own residents on all sources of income no matter where derived, and nonresidents on their income earned from sources within the state.”); id. at 48 n.192 (referring to the power to tax nonresidents on income earned from services actually performed within the state as “[s]ource taxation”).
228 See Changing the Face of College Sports, supra note 64, at 495–98.
229 CAL. REV. & TAX. CODE § 17014(a) (West 2022).
230 Id. § 17016. Arizona’s residency statute mirrors California’s residency statute. See ARIZ. REV. STAT. ANN. § 43-104 (2022).
231 34 TEX. ADMIN. CODE § 3.71(a) (2022).
home in another state." The Texas Administrative Code clarifies that “[a] person ... may be a resident of more than one state at a time.” Comparing these residency statutes, it is possible for an individual to be a resident in both California and Texas at the same time for tax purposes. To counteract the effects of being subject to taxes in multiple states, some states offer a tax credit.

If a student-athlete moves out-of-state for college, they will need to navigate the residency requirements of both their home state and college state to determine in which state they are a resident (or whether they are a resident of both states). If the student-athlete is a resident of both states, they will need to file state tax returns in both states and determine if there are any applicable tax credits in either state. It may be beneficial for a student to become a resident in the state where they attend college. For example, if the college student is from California, where the tax rate is particularly high, and they are attending a college or university in Texas, where there are no state income taxes, it would be beneficial for the student to declare residency in Texas for tax purposes. However, even if the student becomes a Texas resident, California still requires the student to pay part-year taxes for all income earned while a resident of California and all income from a source within California. Calculating part-year taxes can be extremely complicated.

232 Id.
233 Id.
234 See id.; CAL. REV. & TAX. CODE § 17014 (West 2022).
235 See, e.g., CAL. REV. & TAX. CODE § 18001(a) (West 2022) (allowing California residents to take a tax credit “for net income taxes imposed by and paid to another state”); ARIZ. REV. STAT. ANN. § 43-1071 (2022) (offering a tax credit similar to California’s tax credit); Alan Pogroszewski, When is a CPA as Important as Your ERA? A Comprehensive Evaluation and Examination of State Tax Issues on Professional Athletes, 19 MARQ. SPORTS L. REV. 395, 417 (2009) (listing tax credit statutes for select states and indicating Texas’s lack of a tax credit). Although Texas does not have a tax credit, Texas also does not have any state income tax; therefore, Texas residents are not double-taxed if they are residents of Texas and another state. See id.
236 Changing the Face of College Sports, supra note 6464, at 496 (“Depending on how any given state defines ‘residency’ for tax purposes, it is possible that student-athletes may find themselves having dual-residency status.”).
238 See Pogroszewski, supra note 235, at 417.
239 See CAL. REV. & TAX. CODE § 17041(b), (i) (West 2022).
240 For example, in California, nonresident or part-year resident taxpayers must first calculate their taxes as if they were fulltime California residents. See Law Summary: Nonresident or Part-Year Resident Tax Liability 2002 and Subsequent Years, CAL. FRANCHISE TAX BD., http://web.archive.org/web/20161221014249/http://www.ftb.ca.gov/Law/summaries/NonResTxCA_2002S.pdf [http://perma.cc/54NE-ANQE] (last updated Mar. 22, 2010). They must calculate their California tax rate by determining the percentage of taxes they would need to pay over their adjusted gross income (AGI) if they were full...
2. Characterization and Source of Income

The characterization of NIL income has been the issue of tax controversy for some time.\textsuperscript{241} If income is considered compensation for services, then the income will be taxed where the services were rendered.\textsuperscript{242} Conversely, if income is considered royalties, the income will be taxed in the student-athlete’s residential state.\textsuperscript{243} Student-athletes are not earning money for their participation in games;\textsuperscript{244} however, their NIL only has value because of their performance on the field. Because the connection between student-athletes’ performance on the field and their NIL income is indirect, and the NCAA’s reluctance to allow a pay-for-play model, student-athletes’ NIL income will likely be classified as royalties and taxed at the student-athletes’ place of residence.

3. Implications of the Jock Tax

Prior to 1990, professional athletes did not make enough money to be targets for taxation.\textsuperscript{245} In the 1990s, professional athletes’ salaries began to rise significantly, causing the potential revenues from taxing the professional athletes to outweigh the cost of collecting such taxes.\textsuperscript{246} In April of 1992, Philadelphia was in a financial crisis that was estimated to grow to a $1 billion deficit by 1996; therefore, the City had to get creative with raising revenue to decrease this deficit.\textsuperscript{247} As part of these efforts, Philadelphia became the first city to tax professional athletes on income earned within the city.\textsuperscript{248} Many other cities and states have followed Philadelphia’s lead to increase their revenue, but some California residents. See id. They then determine their California AGI which includes “any income from a source within California and income from non-California sources while taxpayer was a resident of California.” Id. To calculate California taxable income, the taxpayer must subtract a prorated deduction amount based on a calculation of their deductions as if they were a permanent California resident. See id. Then, they must subtract the prorated amount from their California AGI. See id. The taxpayer must then multiply their California taxable income by their California tax rate to determine their California tax. See id. Tax credits are then deducted by using a prorated amount of tax credits based on the tax credits that would be allowed if the taxpayer were a full time California resident. See id.

\textsuperscript{241} Whitlock, supra note 140, at 20 (“The characterization of payments for the use of name and likeness rights is not new. It has been the subject of both international tax controversies, and domestic tax controversies.”); see also News & Brews Sports Biz, supra note 98, at 14:10 (explaining the complications of income from social media influencing).

\textsuperscript{242} See Whitlock, supra note 140, at 20.

\textsuperscript{243} See id.

\textsuperscript{244} See NCAA Interim Policy, supra note 80 (clarifying that NIL income is not pay-for-play).


\textsuperscript{246} See id.


\textsuperscript{248} See id.; Fratto, supra note 245, at 40.
states had other motivations. Illinois Senator John Fullerton created a bill, informally titled “Michael Jordan’s Revenge,” that only taxed athletes from states that taxed Illinois athletes. The nonresident income taxes imposed on professional athletes are commonly referred to as “jock taxes.” The jock taxes create an issue of double taxation. Many states that do not impose the jock tax do not give a tax credit for jock taxes paid in other states. Some states have an opportunity for professional athletes to get a tax credit to avoid the issue of double taxation; however, the tax credits do not eliminate the issue of double taxation.

Some states have jock tax statutes specifically directed at professional athletes, but it is likely that student-athletes will also fall under those statutes. Even if student-athletes do not fall under the current jock tax statutes, these statutes may be amended to include student-athletes or new statutes could be created to tax student-athletes in a similar way. Conversely, because student-athletes’ NIL income is not classified as pay-for-play, the student-athletes are not earning income when they are in another state for a game. It is a stretch for forum states to try to tax the student-athletes’ NIL revenue as income earned from a source within the state when the student-athletes are merely

249 See Fratto, supra note 245, at 40 (“Some [statutes] were enacted purely as a method of generating revenue; however, others were enacted in retaliation to recover tax revenue that was lost to other states.”).
250 See Richard E. Green, The Taxing Profession of Major League Baseball: A Comparative Analysis of Nonresident Taxation, 5 SPORTS LAWS. J. 273, 277 (1998). For example, if State X taxes athletes from Illinois, but State Y does not tax Illinois athletes, then Illinois will tax State X athletes, but not State Y athletes. The Chicago Bulls’ basketball players paid approximately $4,000 in taxes to California, and Illinois credited the players for those taxes. See id. Fullerton claims that the purpose of the bill was equity and fairness. Id. at 277 n.24.
251 See Show Me the Money, supra note 9797, at 33 n.103 (using the term “jock tax” to refer to “the trend among taxing authorities toward levying state and local income taxes on traveling business professionals, particularly visiting professional athletes”).
253 See, e.g., DiMascio, supra note 97, at 955 n.23 (“Illinois refuses to grant a credit to its resident athletes for taxes paid to a foreign jurisdiction, effectively double taxing the athletes on dollars allocated to another jurisdiction. It is the only of the 20 states that impose a jock tax that does not grant such a credit.”).
254 See Show Me the Money, supra note 97, at 33 n.104; Changing the Face of College Sports, supra note 64, at 497 n.260.
255 See News & Brews Sports Biz, supra note 98, at 04:12 (“Many states have specific statutes related to tax for professional athletes. College athletes are probably going to fall under those same statutes and therefore be subject to tax in those same jurisdictions.”).
256 See Show Me the Money, supra note 97, at 46–47 (“Should student-athletes be paid in the future, they will be subject to the jock tax in the same fashion that professional athletes currently are.”); see also Wittry, supra note 96 (mentioning that allocation of taxes for student-athletes will be partly “based on professional athlete rules that are already out there”).
257 See generally NCAA Interim Policy, supra note 80 (clarifying that NIL income is not pay-for-play).
there to perform at a game and are not being paid to play in the game. As mentioned in the previous Section, the states can try to argue that the student-athletes’ NIL only has value because of their performance in games, so the income is still derived indirectly from their performance within the forum state. This argument seems circular; yet many arguments related to the concept of amateurism appear to be similarly circular. Additionally, even if states succeed using this argument, the allocation of NIL income to the performance in a specific state would present administrative difficulties.

If student-athletes are required to pay some type of jock taxes, they will need to keep track of how much income is earned while visiting another state. This becomes even more burdensome if a student-athlete remains a resident in the state where they lived before going to college (home state). This student-athlete will still need to pay taxes, at minimum, in their home state and in the state where they attend college (college state). This would require student-athletes to accurately track and allocate their NIL revenue between income earned while in their college state and income earned while in their home state. This is even more burdensome than the jock tax for professional athletes because professional athletes are generally residents in the state where their team is located.

The double taxation and the record keeping requirements of the jock tax are highly onerous for student-athletes. The possible implication of jock taxes for student-athletes further increases the complexity and compliance requirements for student-athletes.

B. Recruiting Inequity

The state tax effects of NIL deals could create a recruiting inequity between different schools. Student-athletes expecting to earn large amounts of NIL income will be enticed to attend institutions in states with little to no income tax. For example, if a top-recruited high school senior football player is deciding between Texas A&M in Texas (with no income tax) and UCLA in California (with high income tax), this football player would have

---

258 See supra Part III.A.2.
259 See Nat’l Collegiate Athletic Ass’n v. Alston, 141 S. Ct. 2141, 2167 (2021) (Kavanaugh, J., concurring) (describing the NCAA’s argument that student-athletes’ compensation must be restricted to conserve the concept of amateurism as “circular and unpersuasive”).
260 See Show Me the Money, supra note 97, at 47.
261 See McCann & Raiola, supra note 124 (“[S]chools in one of the nine U.S. states without a tax on wages (Alaska, Florida, New Hampshire, Nevada, South Dakota, Tennessee, Texas, Washington and Wyoming) could enjoy a potential recruiting advantage on account of their tax laws.”).
262 See id.
much more net profit by choosing Texas A&M over UCLA. The tax differential between states is more important for student-athletes generating large amounts of NIL income; therefore, the states with little to no income tax will have more of an edge in recruiting top athletes. Additionally, the recent introduction of the portal for student-athletes allows them to transfer between schools more easily. The combination of the portal and the tax inequality between states may motivate many students to transfer from schools in states with high tax rates to schools in states with low tax rates. Conversely, the tax inequity between states is just one of many factors student-athletes may consider when choosing a school (or choosing to transfer schools). It is unlikely that the mere tax inequity between states will be a controlling factor in deciding where to attend school.

Ultimately, many federal and state ramifications of the new NIL payments create burdensome compliance difficulties for young and inexperienced taxpayers. An eighteen-year-old is expected to calculate and file taxes correctly while prioritizing academic success and athletic performance. If a student-athlete fails to track, calculate, and pay taxes correctly, there are consequences. If taxes are left unpaid or underpaid, student-athletes will be subject to an additional payment as a penalty. For example, Lawrence Taylor, former linebacker for the New York Giants with appearances in The Waterboy and Any Given Sunday, filed a false tax return and admitted to not paying over $83,000 in taxes. 

---


264 See Eli Boettger, College Basketball’s Most Powerful Force: The Transfer Portal, ATHLETICDIRECTORU, http://athleticdirectoru.com/articles/college-basketball-transfer-portal/#:~:text=In%20October%202018%2C%20the%20NCAA,days%20to%20publicize%20the%20information [http://perma.cc/EU4A-TGEQ] (last visited May 11, 2022) (explaining that the transfer portal, which came into existence in October of 2018, enables student-athletes to transfer schools more easily and allows them to contact other schools without their coach’s permission).

265 See Show Me the Money, supra note 97, at 40–42, 47–49 (analyzing the recruitment inequity that would arise if student-athletes were paid for their participation in athletics).


267 See I.R.C. § 6654.

268 See I.R.C. § 6654(a).

269 See I.R.C. § 7201.
taxes. Taylor was sentenced to ninety days of house arrest. Even though student-athletes are subject to the extra tax compliance burdens outlined throughout this article, it is crucial that they refrain from going out of bounds such that they are subject to extra penalties. Congress, the NCAA, and schools must work together to help student-athletes maintain adequate compliance to stay within bounds.

IV. SOLUTIONS

A. Develop an Online Platform for the Education and Assistance of Student-Athletes

The most helpful and immediate solution to the new tax burdens for student-athletes is to create a digital platform to educate student-athletes about their rights and responsibilities, to help keep track of and make recommendations for their finances, to aid students in taking steps to minimize their tax burdens (including forming an LLC), and to help students with calculation and filing of necessary tax documents. This software could be even broader and incorporate an interface for student-athletes to connect with NIL deals and agents. It should be available to all student-athletes (and recruits) throughout the country. The software should also have many how-to pages and videos on helpful processes, such as explaining what kind of expenses need to be tracked and how to create an LLC. In addition to educational material, the platform should connect student-athletes to lawyers and CPAs who specialize in student athletics. There should be staff working to answer questions through messaging, over the phone, or on an online interface, such as Zoom, for digital meetings. The student-athletes should have the opportunity to create continuing business relationships with specific attorneys and/or CPAs and be able to communicate with them through the app. The software should be available online and should be made into an application for mobile devices. Student-athletes are busy and always traveling for games, so


271 Id.

272 There are already some platforms that are limited to management of sponsorships and education materials on finances and business. See, e.g., NIL Protect Pro, ATHLIANCE, http://athliance.co/nil-protect-pro/ [http://perma.cc/T73X-J22Z] (last visited Aug. 19, 2022). I am suggesting a software that combines these features with the opportunity for students to acquire attorneys and CPAs. See id. Further, the software would be user-friendly, marketed toward young student-athletes, and available both online and as an app. This app would be a one-stop shop to help student-athletes manage the new responsibilities associated with NIL income.
having the software available in an app will allow them to record certain expenses or ask questions with ease.

This platform would be costly because at the start, it would require, at a minimum a CPA, a financial advisor, a specialized sports law attorney, a software developer, and a graphic designer. The software should be a subscription-based online software to offset the large costs. The subscription should have a range of levels for how much assistance a student-athlete would like to use. For example, a student-athlete that merely wants access to the educational materials should not pay nearly as much as a student-athlete who wants to use the app for help from a lawyer to create an LLC and help from a CPA to calculate and file all necessary tax paperwork. It may reach a level where, if a student-athlete wants to utilize services from the CPA and lawyer, the subscription cost is decided on a case-by-case basis.273 The cost of the basic subscription level will be minuscule in comparison to the amount of money potentially saved on taxes and related expenses. This cost could possibly be covered by schools, especially schools with Division I athletic programs that have student-athletes who have the potential to make large amounts of money from their NIL.274 This could even become a recruitment incentive for attending a particular school. Conversely, this recruiting incentive could amplify the unfair competitive advantage of certain schools.275 It may be more beneficial for the NCAA to bear the cost of the software and make it available to all NCAA student-athletes.276

This software will be the most effective solution because it would be available to all student-athletes throughout the United States, would be affordable for student-athletes to begin utilizing, and would not require cooperation from the government, schools, or the NCAA. This solution is not perfect, however, because the creator of this platform would likely be a private company that is willing to fund the creation of the software and has the workforce in place. Further, the software

273 At the base level, for student-athletes making less money from NIL deals that would adequately benefit from a one-size-fits-all tax education, minimization, and compliance package, there would be standard subscription costs. As student-athletes make more money from NIL deals and require more hands-on help—along with a more customized tax minimization and compliance plan—the cost would be determined on a case-by-case basis.

274 Such lawyers must account for professional conduct rules regarding payment from third parties. See MODEL RULES OF PRO. CONDUCT r. 1.8(f) (Am. Bar Ass’n, 2018) (preventing compensation from third parties “unless (1) the client gives informed consent; (2) there is no interference with the lawyer’s independence of professional judgment or with the client-lawyer relationship; and (3) information relating to representation of a client is protected…”).

275 See supra Part III.B.

276 See supra note 274.
would need to address not only federal and state tax requirements, but also contain an algorithm to address the variables associated with any combination of multiple states of residency. Nonetheless, there is a potential to make a large return on investment given the many users that would be paying for access to the proposed software, if it becomes successful.

B. Tax Support Centers Within Athletic Departments

Another possible solution that could alleviate the student-athletes’ complicated tax compliance burdens is for athletic departments to provide resources to assist student-athletes. Such resources could include educational materials, financial advisors, CPAs, and/or specialized attorneys. Athletic departments should provide these resources to student-athletes to help address the complications of their tax returns. Athletic departments could provide legal counsel and/or tax specialists to help student-athletes structure their finances for tax minimization and to help calculate their taxes for their estimated quarterly payments and annual tax returns. The school could have in-house counsel dedicated to student-athletes, or the school could contract with a team of independent consultants, giving students more flexibility in what kind of help they desire.

Alternatively, the NCAA could have a centralized tax support center available to all NCAA student-athletes. In this scenario, the lawyers would work for the NCAA, and student-athletes would have access to them virtually. This option is less ideal though, as the lawyers would be less accessible than if they were in each school.

In either situation, the attorney’s role would be to counsel student-athletes on their burdens and liabilities regarding taxes from NIL income; help student-athletes create an LLC, and counsel them on their responsibilities as members of the LLC, including keeping track of expenses; and/or be an agent for student-athletes in signing NIL deals and managing LLCs. The CPA would help with tracking of expenses, calculating quarterly estimated payments, and filing tax returns.

It is likely that the assistance, if paid for by the NCAA or athletic departments, could also be excluded. As mentioned above, the IRS has traditionally opined that student-athletes’ scholarships qualify for the exclusion. Additionally, the Court has interpreted that, as long as scholarships and grant funds are not received in exchange for services rendered, the scholarship or
grant qualifies for the exclusion.\textsuperscript{277} This service would not be compensation for the student-athletes’ participation in sports, so it would likely fall under the scholarship tax exclusion and would not add another tax obligation for the student-athletes.

This type of assistance would be costly. It would not be a good solution for all schools. This type of tax support would be more fitting for Division I schools that have ample income from athletics and are bringing in many student-athletes receiving large compensation from their NIL. Further, although this would be costly for the institutions involved, universities have an incentive to prevent student-athletes from dropping the ball on their tax liability. Universities, understandably, want to avoid a headline that a student-athlete at their university was indicted for tax evasion. These services could even be used as a recruiting incentive to compel athletes to attend that school. Providing legal and tax help to student-athletes is a highly valued incentive that would otherwise cost these individuals a potentially large sum. This is not a one-size-fits-all solution, as it will only be feasible for a handful of schools. It is a great option for those schools, but not enough to alleviate the burden of the student-athletes across the country being thrown into complex tax returns with little to no tax experience.

An additional barrier to implementation of this solution is that this type of support center is not currently allowed by the NCAA guidelines.\textsuperscript{278} This tax support center is not related to the student-athletes’ education; therefore, the new guidelines from NCAA v. Alston are not broad enough to provide this type of support center.\textsuperscript{279} The NCAA is generally only concerned with “protect[ing], support[ing] and enhanc[ing] the physical and mental health and safety of student-athletes.”\textsuperscript{280} However, the student-athletes need to be able to satisfy their tax obligations in order to focus on their physical and educational well-being.\textsuperscript{281} The

\textsuperscript{277} See, e.g., I.R.C. § 117; Rev. Rul. 77-263, 1977-2 C.B. 47; Letter from John A. Koskinen, Internal Revenue Serv. Comm'r, Internal Revenue Serv., to Hon. Richard Burr, United States Sen. (June 27, 2014), http://www.irs.gov/pub/irs-wd/14-0016.pdf [http://perma.cc/35YHCYSH] ("It has long been the position of the Internal Revenue Service that athletic scholarships can qualify for exclusion from income under Section 117."); Smith v. Comm'r, 51 T.C.M. (CCH) 1348 (1986) (holding that funds received by a graduate assistant were excludible because they were not compensation for her services rendered). But see Bingler v. Johnson, 394 U.S. 741, 755–58 (1969) (holding that a Ph.D. student’s stipend did not qualify for the Section 117 exclusion because it was compensation for services, rather than a scholarship).

\textsuperscript{278} See NCAA Division I Manual, supra note 33, at 218.


\textsuperscript{280} See NCAA Division I Manual, supra note 33, at 2.

\textsuperscript{281} Id.
NCAA should promulgate guidelines allowing for these tax support centers within universities.

C. A College Course on Student-Athlete Tax Liabilities

Another possible solution is for colleges to require that student-athletes participate in a class about their tax responsibilities. “Give a man a fish, and he will eat for a day. Teach a man to fish and he will eat for a lifetime.”282 Some students begin college without even knowing whether they need to pay taxes or how that process begins. Some students may not need to file taxes, but student-athletes making money from their NIL will need to learn quickly. In their first term of college, student-athletes should be required to take a semester-long class (or if the school uses trimesters, then a trimester-long class) taught by tax lawyers and/or CPAs where they are taught and tested on tax planning techniques and calculation of taxes for their NIL income. Although an entire course could be described as overly onerous, the skills the student-athletes learn in this course will be utilized throughout the remainder of their lives. The course could be as little as one course unit/one lecture per week, throughout the term.

In addition to the semester-long class, athletic departments should provide references for students who need more help, including information for local CPAs and tax lawyers and handouts for useful information. The school should ensure the information is frequently updated and available to students online and/or in athletic departments.

Teaching student-athletes how to handle their own taxes will not only help them with their tax burdens in college, but it will also prepare them to handle their taxes for life. However, it is unlikely that one class will be sufficient to give students the tools to be able to calculate their own taxes. This solution is fairly weak because it leaves the burden of financial planning on the student-athletes, who are already full-time students and athletes. In addition, this would require schools to enforce the requirement for students to take the class. It is unlikely that all schools will cooperate in creating and offering this course without being required to do so. Schools could be required to enforce this type of education either through an NCAA policy or federal legislation. An NCAA policy would be more ideal because it would be quicker to enact and more cost-effective than federal action.

CONCLUSION

With NCAA v. Alston, new state legislation, new NCAA guidance, and potentially new federal legislation, the landscape of financial planning and taxes for student-athletes is turning into a new game entirely. These positive changes have added new burdens to student-athletes’ tax responsibilities, including higher tax payments, more complex calculations, and more record keeping requirements.

As independent contractors, student-athletes need to pay twice as many employment taxes as employees and will need to pay quarterly estimated payments. They will benefit from being able to deduct certain expenses as ordinary and necessary business expenses; however, this will require thorough record keeping and a determination of the student-athletes’ tax home. Student-athletes could create an LLC to minimize taxes, but this creates more responsibilities for a student-athlete. Additionally, for purposes of state taxes, student-athletes must determine the state(s) in which they will be required to file state taxes (which could be multiple states), based on their states of residency, the characterization of their income, and the implication of the jock tax. Because of student-athletes’ young age and lack of experience with taxes, and in light of the complicated tax compliance burden as a result of their newly allowed NIL income, solutions must be implemented to help student-athletes with tax compliance to avoid a host of potential penalties.

The most effective solution is to create a digital platform to educate student-athletes, help keep track of and make recommendations for their finances, aid students in taking steps to minimize their tax burdens (including forming an LLC), and help students with calculating and filing necessary tax documents. Tax support centers staffed with lawyers and accountants, in addition to requiring a course on tax and finance responsibilities, could also help student-athletes avoid penalties for tax compliance errors. By providing student-athletes with the tools, education, and advice they need to legally minimize their taxes and stay on top of filing and payment requirements, we can better protect them from dropping the ball on the compliance of their tax returns.

283 See Taxing College Athletes After NCAA v. Alston, TAX NOTES TALK, at 17:07 (July 16, 2021), http://open.spotify.com/episode/0LeU6JZQIT13trB6viGmvPv?si=84C2Rw9iQ NCTfbhWwTg0OA (“[Taxation of student-athletes] remains the wild, wild west.”).
284 See supra Part II.B.1, 3–4.
285 See supra Part II.B.2.
286 See supra Part II.C.
287 See supra Part IV.A.
288 See supra Part IV.B–C.
The Chapman Law Review is published by its student members at Chapman University Dale E. Fowler School of Law. The Chapman Law Review can be reached online at www.chapmanlawreview.com or by e-mail at chapman.law.review@gmail.com. Submission inquiries can be sent to chapmanlawreview.submissions@gmail.com. The office of the Chapman Law Review is located in Donald P. Kennedy Hall on the campus of Chapman University, One University Drive, Orange, CA 92866.

The views expressed in the Chapman Law Review are solely those of the authors and in no way reflect the views of the Chapman Law Review, Chapman University Dale E. Fowler School of Law, or Chapman University.