Psychological versus economic models of bounded rationality

Don Ross

University of Cape Town and Georgia State University
don.ross@uct.ac.za

Abstract

That the rationality of individual people is 'bounded' – that is, finite in scope and representational reach, and constrained by the opportunity cost of time – cannot reasonably be controversial as an empirical matter. In this context, the paper addresses the question as to why, if economics is an empirical science, economists introduce bounds on the rationality of agents in their models only grudgingly and partially. The answer defended in the paper is that most economists are interested primarily in markets and only secondarily in the dynamics of individual decisions – specifically, they are interested in these dynamics mainly insofar as they might systematically influence the most useful approaches to modeling interesting markets. In market contexts, bounds on rationality are typically generated by institutional and informational properties specific to the market in question, which arise and are sustained by structural dynamics that do not originate in or reduce to individuals' decisions or psychological dispositions. To be sure, these influences *interact* with psychological dispositions, so economists have reason to attend to the psychology of valuation. But no *general* model of bounded rationality should ever be expected to feature in the economist's toolkit, regardless of the extent to which psychologists successfully identify specific human cognitive limitations. Use of moderate rational expectations assumptions should be understood in this light. Such assumptions are readily relaxed in specific applications, and in ways customized to modeling circumstances, that modelers, experimentalists and econometricians are making steadily more sophisticated.

Keywords

bounded rationality, rational expectations, revealed preference theory, neo-Samuelsonian methodology

JEL classification codes

A12, B20, B40, D01, D03

1. Introduction[1]

Debates over whether economic models should feature bounded or boundlessly rational agents are typically presented as decidable by empirical facts about people making decisions. But any debate decided in that way should be over as soon as it

[1] I thank the guest editors of this special issue for most helpful criticisms of an earlier draft.

starts – people are *obviously* not boundlessly rational agents, since nothing is or could be such an agent. Understanding controversies about bounded rationality (BR) in this way involves confusing economics with psychology. Other debates over BR revolve around the idea that economists are participants in a conceptual investigation of a general set of normative ideals associated with rationality. But these discussions confuse economics with formal decision theory.

The present paper is motivated by, though it makes no pretense to comprehensively answer, the following question. Why, if economics is an empirical science, do economists introduce bounds on the rationality of agents in their models only cautiously and partially? Summarized broadly, my answer will be that most economists are interested primarily in markets and only secondarily in the dynamics of individual decisions – specifically, they are interested in these dynamics mainly insofar as they might systematically influence the most useful approaches to modeling interesting markets. In market contexts, bounds on rationality are typically generated by institutional and informational properties specific to the market in question, which arise and are sustained by structural dynamics that do not originate in or reduce to individuals' decisions or psychological dispositions. To be sure, these influences *interact* with psychological dispositions, so economists have reason to attend to the psychology of valuation. But no *general* model of bounded rationality should ever be expected to feature in the economist's toolkit, regardless of the extent to which psychologists successfully identify specific human cognitive limitations.

2. Should bounded rationality be empirically contentious?

BR in its most general sense refers to a model of decision or choice in which it is not supposed that the choosing agent draws all logically and mathematically valid consequences of all physically accessible information that is relevant to determining whether the agent's choices are optimal for her. 'Physically accessible' can be unpacked as 'meaning within spacetime reach, and detectable and discriminable by the agent's input detectors'. 'Relevant' means 'bears upon the expected utility of an action in the agent's choice set'. Given these definitions, non-economists might be expected to wonder how or why it can be, or ever could have been, controversial in economics that actual choices should be modeled as boundedly rational. After all, any actual choice must emerge from physical information processing, and all information-processing episodes involve finite time and computational resources. Did economists, at least before Herbert Simon came along, really imagine that economic choices are made by boundlessly rational agents that transcend physical finitude?

Behavioralist critics of mainstream or 'neoclassical' economics[2] often say or imply that economists have traditionally modeled choice problems as these would arise

[2] I use the word 'behavioralist' here in light of the fact that critical behavioral economists who think that mainstream economic theory requires fundamental revision in light of experimental evidence concerning BR are not behaviorists.

for a being that is omniscient with respect to the past and present, but not with respect to the future (because it can be exogenously shocked). Recent literature in economic methodology, philosophy of economics, and empirical behavioral economics abounds with examples of writers who view such modeling as self-evidently misguided. Thaler (1992) set the template for such criticism. Whereas Ariely (2008) works a populist vein that is inattentive to actual views of historical economists, Camerer, Loewenstein and Prelec (2005) press the anti-establishment point in an article published in an AEA journal. Some writers distinguish a few great economists who, it is suggested, deserve congratulations for having avoided the foolishness of modeling boundless rationality; Keynes often attracts such praise, as does Knight, and occasionally Marshall. The hero is the originator of BR as an explicit idea, Herbert Simon (1957).

The behavioralist critique is often assimilated to arguments against the use of macroeconomic models based on the rational expectations hypothesis (REH). Behavioralists typically argue that, as an empirical matter of fact, people do not make incentivized choices as if they had rational expectations. Therefore, this style of argument concludes, such expectations should give way in models to expectations that can be generated by boundedly rational processes.

The REH originated in a microeconomic context, and was explicitly presented by its originator, Muth (1961), as an alternative to a specific expression of BR, namely cobweb models of updating of expectations about prices. Nerlove (1958) had modeled farmers as operating a heuristic to the effect that next year's prices are predicted by extrapolation from previous years' prices. Muth objected that this rested on an unjustified assumption that farmers could not understand a relatively uncomplicated and more accurate model based on supply and demand trend analysis. The contrasting methodological assumption defended by Muth is not that famers are epistemic gods – i.e., are boundlessly rational. Rather, Muth argues that if the economist can identify a model that makes more accurate predictions by taking fuller account of available information than the model attributed to the agents, then the economist owes an explanation of why the agents can't also identify the superior model. This argument is generalized by Ragot (2012, pp. 187-188) as follows: "If the economist builds a theory or model in which the agents fail to do something that it is in their interests to do, then the economist must justify why they did not do it."

This reasonable – original – version of the REH can be contrasted with Lucas's (1976) stronger principle for application to macroeconomics. Lucas's version is targeted at policy-makers rather than model-builders per se, and enjoins that monetary or fiscal authorities should not operate any policy that, to have its intended effects, depends on the agents whose behavior is to be influenced by the

---

Indeed, they tend to explicitly reject the behaviorist aspects of mainstream theory that accompany the idea of revealed preference. See Ross (2014, Chapter 5) for extended commentary on the historical and logical relationships between behavioralism and behaviorism.

policy gathering or using less information than the policy authorities themselves. This strong REH sometimes encourages the view that macroeconomic policy is futile, as the agents whose behavior is supposed to be altered by such policy will bring about whatever good a policy maker restricted by the REH might try to accomplish anyway, leaving *unintended* policy effects, which are more likely to be harmful than helpful, as the only possible ones.

The futility hypothesis based on the Lucas critique is doubtful. Among other problems it ignores dynamical coincidences and effects of heterogeneous time preferences. For example, the Federal Reserve policy of quantitative easing (QE) that is in place as I write in mid-2013 was designed to stimulate equity markets for awhile, not forever. It seems likely that it has achieved this limited aim, notwithstanding the fact that investors anticipate asset price correction after QE ends. Let us accept for the sake of argument, writing while QE is still in force in September 2013, that equity prices will 'ultimately' return to equilibrium levels that would have prevailed had QE never been adopted. The world might nevertheless be better off, even in the long run, because equity prices were temporarily boosted during an interval when investors weren't sure that the European Central Bank would forestall a collapse of the Euro, or, later, whether US manufacturing and construction would begin to recover before Chinese growth began to slow. That is to say: historians of markets might later conclude that QE prevented a double-dip recession outside Europe thanks to its timing. This example, if events vindicate it, refutes the policy futility conclusion but not the REH. The monetary authority did not need to know anything that investors didn't; it needed only to share in the common knowledge that the summer of 2012 was a dangerous moment for financial markets and that it could be made less dangerous if general softening of demand for equities could at least be deferred for a time.

Debates about the place of BR in economics should thus be separated from debates over the validity of the Lucas critique. Furthermore, acceptance of the point that all actual economic choices are finite episodes that draw on finite information processing resources and finite content should not be taken as implying that purely normative economic analyses are of no empirical relevance. Identifying allocations that are long-run optimal may be highly relevant to welfare analyses, even if policy choices that are guided by knowledge of such first-best solutions would require mechanism designs that influence the incentives of agents who cannot or do not take account of all relevant information or who are prone to statistical errors. Two types of scenarios can be considered here. In a Type 1 scenario, agents understand and accept the analysis that identifies the first-best solution, but because of limitations in their cognitive capacities cannot reliably execute the choices that are entailed by this solution. Such limitations are familiar to the mechanism design literature as 'trembles'. Trembles are equally relevant to situations in which policy-makers do not aim at first-best solutions because these are blocked by strategic entanglements among agents, but agents fail even to converge on vectors of strategies that implement second-best equilibria. The microeconomic literature has produced a range of technologies for modeling such situations, the most important of which are quantal response equilibrium (QRE) analysis (McKelvey & Palfrey 1995, 1998) and global games (Carlsson & van Damme 1993, Morris & Shin 2003). In a Type 2 scenario, cognitive limitations such as inability to resist hyperbolic temporal discounting prevent agents from agreeing with the analyst as to what constitutes the first-best solution.[3] The possibility of Type 2 scenarios gives rise to suggestions for 'nudge paternalism' (Camerer et al 2003, Sunstein & Thaler 2003a, 2003b). Paternalism is 'nudge', as contrasted with unqualified paternalism, only if the analyst can persuasively argue that the agents she is modeling *would* endorse her first-best solution *if* their relevant cognitive limitations were surmounted (Bernheim & Rangel 2008).

There is an additional motivation for modeling unbounded rationality that is independent of an analyst's first-order concern with welfare. This is that agents are interested in their own welfare (so the analyst's interest in it is second-order) and are typically capable of learning from experience. Of course, learning is itself a finite cognitive process so is also subject to bounds. But a ubiquitous form of learning, at least among social animals and certainly among people, is based on observing and copying the behavior of successful agents. Copying behavior without understanding or modeling the strategy that guides that behavior will tend to lead to out-of-sample mistakes; but how serious a problem this is likely to be for an agent depends on the stability of the problem environment. In highly stable environments, or where agents are good at inferring true models from observed behavioral patterns, the whole population might learn, with some lag, whatever is learned by the most cognitively flexible agents (that is, the agents whose capacity for extracting information objectively present in data is least restricted). Of course, it might not; identifying conditions under which copying is efficient at the population level is a target for empirical economics guided by theoretical modeling. But when, as often, we lack a theory of the most cognitively flexible agent in the modeled population – perhaps because this agent should be expected on probabilistic grounds to be at least as cognitively flexible as the analyst – there may seem no specific BR model that should be preferred as an approximation to a model in which agents suffer from no bounds on their learning at all. That is, the analyst in such circumstance models the population as capable of learning everything there is to be learned because the analyst knows only that she cannot specify where limitations arise. This is not a point of arcane theory; such a modeling strategy is often a wise approach for an investor in a large asset market.

---

[3] Mallard (2011, p. 691) offers hyperbolic discounting, in cases where it leads to intertemporal preference revcersals, as an example of a behavioralist restriction on economic optimization that does not involve BR. I disagree. As Ainslie (1992) argues, boundlessly rational hyperbolic discounters should 'bundle' their choices instead of treating them as independent. Failure to bundle by such agents is modeled by economists either as a kind of trembling (e.g. Gul and Pesendorfer 2001) or as failure to accurately compute optima (e.g. O'Donoghue and Rabin 2001). Ross (2010) discusses these issues in some detail.

In light of these considerations, we can distinguish what should be regarded as incontestable *local* applications of BR from an ambitious methodological programme to develop a general BR economics. BR is incontestable when we know, empirically, that the most cognitively flexible agent in a learning population is less cognitively flexible than the analyst, and also have a reliable model of the nature and extent of this flexibility gap. To begin with an easy example, an economic model of foraging wasps can incorporate the assumption that the wasps will not learn to fly under visually transparent barriers, or learn to concentrate on nutritious food they cannot taste, even if the environment is configured in such a way that these behaviors would clearly be part of strategies that would optimize the wasps' survival prospects or reproductive fitness. Where people are concerned, local BR assumptions can sometimes be relied upon if circumstances are novel and allow no scope for learning, or if we know that specific information known to the analyst to be crucial to inferring the true model is unavailable to the agents.

By contrast, an ambitious BR programme is based on identifying systematic limitations to cognitive flexibility which, it is argued, cognitively typical people will exhibit unless specific structures or policies are put in place by economic engineers or 'choice architects' who transcend those limitations. One way of conceptualizing this, in light of the foregoing remarks, is as follows. The reasonable REH as formulated by Ragot imagines the economist giving a model-specific justification as to why some particular agent or agents are modeled as failing to do what is in their interests. An ambitious BR program would identify motivations of this kind that are *not* situation-specific and apply to people in general, or at least to some relatively general class of people across widespread sets of real environments and / or general types of choice contexts.

In the sections to follow I will argue that the psychology of individual decision making amounts to a kind of ambitious BR program for socially isolated human agents, and is well motivated as such. However, I will argue, there is no similarly well motivated ambitious BR program for social science; and as economics is a social science there is no well motivated ambitious BR program for economics. The details of the dependencies claimed here warrant emphasis. Much of the discussion that follows will be concerned with features of markets, as economists distinctively model them. However, the impossibility of a general BR model arises not from factors unique to market contexts, but from situationally variable ways in which social learning transcends psychological limitations. So the point applies to sociology, for example, as much as it does to economics. This is compatible with supposing that economists have reasons, grounded in the narrow range of purposes that drive agents in markets and the relatively transparent institutional rules governing market exchanges, for being less *troubled* by the impossibility of a general BR program than sociologists might be.

3. Bounded rationality in psychology and behavioral economics

The critical behavioralist tradition in economics is diffuse. Davis (2010) divides it into three sub-traditions. The first, pioneered by Kahneman, Tversky and their collaborators (Kahneman *et al* 1982, Kahneman 2011), emphasizes systematic biases in the representation of expected utility. The second sub-tradition is the 'frugal heuristics' school of researchers led by Gerd Gigerenzer (Gigerenzer et al 1999), which aims to identify ranges of problems in which the use of simple heuristic rules outperforms statistical analyses across relatively subtle variations in the paramaterizations of the problems in question. Finally there is the family of *constructed preference* models, as gathered in Litchtenstein and Slovic (2006), which drops the assumption of exogenous preferences, and instead represents preferences as evolving, path-dependent responses to cross-frame consequences of choices made within frames.

At the beginning of Section 2 above I defined BR by reference to choice procedures that fail to gather or compute the implications of in-principle available information that is potentially relevant to *ex ante* optimization. The sub-traditions of behavioral economics distinguished by Davis stand in differing relationships to models of optimization. The constructed preference approach can, and often does, ignore such models altogether. In this respect it most closely resembles practice in the psychology of valuation disconnected from economics. Frugal heuristics researchers, on the other hand, promote their models on the basis of the claim that they are more efficient and reliable in guiding optimal action, within their special domains, than cross-domain optimization models. In other words, frugal heuristics modelers explore circumstances in which boundless rationality is not rationally recommended. It is important to distinguish, as authors in this sub-tradition often do not, between reasons for ignoring information that can be modeled as *ex ante* optimizing in conventional cost-benefit terms, and motivations an agent might have for failing to conform her behavior to one or more axioms commonly used to formally characterize economic agents, such as the weak axiom of revealed preference (WARP). The latter type of adjustments are often rhetorically associated with BR. Manzini and Mariotti (2007) provide an example of such rhetoric in comments accompanying their model of agents in whom a plurality of motivations, including concerns about consumption and effort scheduling, induces restricted violations of the transitivity axiom of WARP. As in other frugal heuristics models. Manzini and Mariotti's agents *do* optimize utility *ex ante*, given their multi-criterial utility functions that range over plans and not merely over separable actions. In Section 4 I will discuss relationships between models of this type and BR as I have defined it here. For now, let us note that such models can at most implicate BR, as I have defined it, in a derivative sense, where axiom violations imply failing to gather or to compute the implications of some in-principle available relevant information.

The Kahneman-Tversky sub-tradition has the most nuanced relationship to standard optimization as modeled using expected utility theory (EUT). The original version of prospect theory, proposed in Kahneman and Tversky (1979), incorporated a straightforwardly psychological modeling component, namely, a cognitive 'editing' process that was postulated as preceding valuation of alternative reward prospects. However, in the absence of a restriction on editing outcomes that would have been ad hoc from the psychological point of view, original prospect theory was incompatible with general EU maximization because it allowed for violation of the dominance axiom of EUT given three or more possible choice outcomes. Cumulative prospect theory (CPT), as offered in Tversky and Kahneman (1992), avoids such violations. This allows it, in principle, to be tested directly against EUT models in estimation of data sets constructed by asking experimental subjects to choose preferred lotteries in which both expected value and outcome variance are systematically and independently manipulated by the experimenter (Harrison and Rutström 2008). Such methods also allow comparison of the empirical performance of CPT against rank-dependent utility (RDU) models (Quiggin 1982; Yaari 1987) in which agents choose as if they become more pessimistic about probabilities as the outcomes associated with the probabilities ascend in their preference orderings.

In light of this, it is tempting to suppose that the move from original prospect theory to CPT unifies BR modeling with standard economics, and turns the relationship between boundless and bounded rationality models into a straightforwardly empirical one. That is, with Tversky and Kahneman (1992) we seem to find psychological methods harnessed to the economist's concern with optimization, thus unifying the two disciplines but fundamentally transforming economic theory in the process through the displacement of EUT. However, this supposition would be too simple, for two reasons. First, the narrative is naïve concerning the empirical status of CPT. Second, it rests on exaggerating the dependence of standard economic theory on EUT.

Explaining these two points requires some review of constraints on identifying different utility models in experimental choice data. Preference structures as modeled using EUT and RDU are independent of points on wealth gradients from which alternatives are evaluated. CPT essentially incorporates RDU and then adds the hypothesis of loss aversion relative to a reference point. Loss aversion can hold either in utilities or probabilities[4] or both. The reason CPT avoids dominance violations is that a specific CPT model combines two RDU functions, one in the gain frame and one in the loss frame relative to the reference point, which is normalized as the zero value. The modeler then assumes semi-differentiability around zero and compares the right derivative at zero to the left derivative at zero.[5] The CPT utility function is constructed by cumulatively iterating one standard RDU model from the top down in the gain frame and another RDU model from the bottom up in the loss frame. This procedure typically makes the overall utility function extremely

---

[4] For example, a decision maker might base decision weight on accurately estimated probabilities where prospects of gains are concerned, but implicitly assign decision weights that reflect pessimism with respect to stated probabilities with respect to loss prospects.

[5] As Abdellaoui et al (2007, p. 1662) document, definitions of loss aversion in the experimental literature have not always been consistent with this approach. However, I suggest that as formal treatments of loss aversion have accumulated, there has been convergence to the procedural definition indicated.

sensitive to the location of the reference point (Wakker 2010). In experiments it is often assumed that the reference point is simply the subject's initial endowment before any incentivized choices are made, and that this point is thus as straightforwardly observable and experimentally controlled as are the dependent variables in EUT and RDU models.

The reliability of this assumption is questionable. As noted above, the general loss aversion hypothesis is ambiguous as between utility loss aversion and probabilistic loss aversion. In a specific application of utility loss aversion we might have

$$U(x) = u(x) \text{ if } x \geq \$0$$

$$U(x) = -\lambda u(-x) \text{ if } x < \$0$$

$$\lambda > 1$$

with reference point $0. But it is open to us to model the agent as assigning different probability weightings to gains and losses, i.e., $\omega(p) = p^{\gamma+}$ for gains and $\omega(p) = p^{\gamma-}$ for losses. Whenever differential decision weighting for gains and losses does not hold, then all differences in risk preferences as between gains and losses must be represented by λ. Experimenters are not allowed to cause subjects to experience real losses relative to their pre-experimental status quo endowments. Thus in experiments that aim to test CPT against EUT or RDU or both, one must postulate $-\lambda$ or $\omega(p) = p^{\gamma-}$ in what is held to be a loss *frame*, that is, in a real *gain* domain that one thinks one has induced the subject to represent as a loss frame. For example, the subject might be rewarded for some achievement task or tolerance of an aversive experience, in hopes that she will represent the endowment as earned, and therefore represent it as then subsequently lost if she forfeits it based on the outcome of a risky choice. Given the noted extreme sensitivity of overall utility function estimations to reference point identification, this resort to hopeful hypothesis in place of experimental control is a matter for serious methodological concern.

Overcoming this concern would require application of psychological probes that are not compatible with the economist's standard revealed preference theory (RPT) conception of preferences as summaries of consistent choices. In fact such probes are absent from CPT methodology as so far manifest. Harrison and Swarthout (2013) find that only two published empirical studies (Rieskamp 2008; Nilsson et al 2011) to date have tested CPT against alternative utility models by offering subjects real, salient incentives over both gain and loss frames, and let the data speak to the question of whether apparently loss averse responses might stem from utility loss aversion or probability loss aversion, as opposed to imposing the former assumption. Both studies estimated λ to be equal to or approximately equal to 1.[6]

---

[6] Rieskamp's estimation involved the constraint that $\lambda \geq 1$.

Thus studies that claim to have demonstrated that CPT models human choices more accurately than EUT have at best shown the need to allow for some rank dependence in the utility functions of at least some agents.

Behavioral economists, and many philosophers (e.g. Hausman 2011) are likely to object at this point that RPT should *not* be maintained as the relevant conception of preference. But this begs the question as to whether CPT unifies psychological and economic models of valuation, by simply asserting that the economic models should give way. Of course the question of whether methodologies that incorporate RPT are worth deploying can be independently addressed. Hausman's philosophical arguments against RPT, for example, are unrelated to empirical efforts to confirm CPT. If use of RPT were considered to vitiate the soundness of *any* methodology, then one could accept Harrison and Swarthout's criticism but turn its general implication upside down. It is indeed the case, someone following Hausman might say, that none of the experiments Harrison and Swarthout survey have the implications claimed for them, but this is because these experiments rely on the methods of economics. Such a critic could go on to conclude that CPT should be abandoned in favor of original 1979 prospect theory and that work should focus on developing psychological probes – perhaps with help from fMRI machines – that could provide independent evidence of loss framing.

But there are reasons why most economists do not think they would be best advised to simply turn into psychologists. We will now review some of those reasons, and then ask what difference they make to models of BR.

4. Bounded rationality in conventional economics

Above, we imagined seeking a psychological probe that would enable us to determine, independently of behavioral evidence, whether experimental subjects frame endowments gained and then lost in the lab as losses. Suppose that we develop such a probe at some point. Do we expect its use to reveal that there is a psychological generalization about this framing – for example that all people, or all men, or all risk averse people, or even all sober men over the age of 25 when faced with choices over house money, frame in the way suggested? This would reassure us that we can use behavioral tests – e.g. lottery choices – to distinguish amongst CPT, EUT and RDU models. Of course, we would still also need to behaviorally distinguish between utility loss aversion and probability loss aversion, and we would still be in the position of looking for our first empirical evidence that what causes CPT to out-perform EUT, when it does, is more than just the rank-dependence of a CPT preference profile. Still, we would have addressed the main methodological worry raised in Section 3.

However, the conjecture that a strong psychological generalization about loss framing might be confirmed is a very bold one, and in conflict with evidence that such framing can readily be manipulated by subtle changes in experimental design (Plott and Zeiler 2005, 2007). When it comes to cognitive dispositions to respond to changes in monetary incentives, psychologists and economists alike have learned to expect heterogeneity (Bardsley et al 2010). Furthermore, economists are typically interested in aggregate responses to incentive changes by reference groups, not in the behavioral adjustments of individuals. Since institutional variables almost certainly influence most framing decisions, we would not expect psychological framing dynamics to be independent of social interactions. Thus the more modest and plausible hypothesis is that loss framing – like most cognitive framing by people – is controlled by the interaction of psychological dispositions (which are themselves likely stochastic and also sensitive to individual learning histories) and social and environmental cues, structures and signals.

The relevance of framing to BR as defined in Section 2 is direct: framing prior to choice amounts to an *a priori* restriction on kinds of information that will be taken into account. So if prospects for a general theory of choice framing were promising, so would be prospects for a general theory of BR.

The pursuit of general theories of some *elements* of framing is the very lifeblood of contemporary psychology. Generalizations about human framing of perception and of cognitive and emotional attention are routinely offered, tested, criticized and refined by psychologists. This kind of progress is mainly achieved by isolating individual research subjects in artificially simplified laboratory environments where experimenters can carefully control restricted ranges of available framing influences. Where choice is concerned, such methods are directly useful for studying relatively stimulus-bound decisions that individuals generally make with relatively little attention from other people or from institutions; for example, whether to put strawberry or blueberry jam on one's toast (assuming that both were already purchased and in the fridge).

I will argue that these methods of psychologists are not at all likely to yield a general theory of the framing of those choices that *are* socially and institutionally mediated – and these are the kinds of choices that mainly interest economists, along with other social scientists. Before we get to that, however, it is important to note the profound effect that five decades of the flourishing cognitive science research programme have had on the methodological respectability of standard economics. Cognitive science was launched through the successful revolt against behaviorism in psychology and linguistics. As Hausman (2011) argues, this makes economists' reliance on RPT, with its emphasis on observed behavior and its disavowal of latent influences on choice, appear to represent methodological atavism. Less philosophically sophisticated versions of Hausman's argument are stock in trade for behavioralist critics of mainstream economics who promote a new basic theory anchored in BR.

The key to appreciating why BR should be *of interest to* economists but should not be expected to fundamentally transform economic theory lies in seeing why economists' reasons for sticking with RPT are not vulnerable to the considerations that led psychologists to abandon strict behaviorism. I here synopsize a perspective that is extensively developed in Ross (2014) under the label of *neo-Samuelsonian* methodology.[7] Theorists and practitioners identified with the neo-Samuelsonian methodology include Vernon Smith (see his 2008), Ken Binmore (2009), Charles Plott (1986), and the authors gathered in Cox and Harrison (2008). It is Samuelsonian because it identifies preferences with statistical patterns in observed choices. It is *neo*-Samuelsonian because it allows that the processes that generate choice data typically *include* latent psychological components. (Behavioralists, by contrast, *identify* choice processes with such latent causal dynamics; see Glimcher 2011 for the best developed example and Ross 2012 for critical context.)

The apparent tension in endorsing a version of RPT while also acknowledging the relevance to economics of latent psychological processes arises from a distortion in the interpretation of the history of mid-20th century economics that has become extremely common due to two complementary influences. The first of these is the pedagogy of introductory microeconomics textbooks, which often begin with idealized individual consumers and then aggregate them into frictionless markets. The second is informal association, not grounded in any formal or empirical theory, of rational decision theory with individual choice. Both of these aspects of economics pedagogy sit uncomfortably with the fact that in the basic classic text of neoclassical economics, Samuelson's *Foundations of Economic Analysis* (1947), no individual economic agents, rational or otherwise, put in an appearance. The sole mention in the book of *Homo economicus* is an ironic reference to his role in the arguments of others. Samuelson does not leave us to speculate about why *H. economicus* is absent. He concludes his chapter on the basic mathematics of RPT with these words:

> Many writers have held the utility analysis to be an integral and important part of economic theory. Some have even sought to employ its applicability as a test criterion by which economics might be separated from the other social sciences. Nevertheless, I wonder how much economic theory would be changed if … [the conditions identified in the utility analysis] … were found to be empirically untrue. I suspect, very little. (Samuelson 1947, p. 117)

The Weak Axiom of Revealed Preference (WARP) was in no sense regarded by Samuelson as a proto-psychological construct. As of 1938, when he proposed the axiom under a different name, he was urging the elimination of the concepts of utility and preference from economics. WARP was instead regarded by Samuelson as a measurement-theoretic specification of some restrictions on demand that allow for uniqueness of equilibrium solutions to models of markets. Also necessary for this objective is symmetry of the Slutsky matrix of substitution items in a demand schedule. Houthakker (1950) extended the analysis to apply not just to choices between pairs of bundles, as in WARP, but to choices over chains of bundles,

---

[7] This discussion accords with that of Hands (2013), who recognizes the coalescing methodological school I call neo-Samuelsonian under the label "contemporary revealed preference theory".

yielding the Strong Axiom of Revealed Preference (SARP). SARP implies Slutsky symmetry, and axiomatizes the strongest possible relationship between ordinal utility theory (ORT) and WARP: given Slutsky symmetry ORT and WARP are *equivalent* – one applies exactly and only when the other does. Then any equilibrium in an economic choice problem involving consumption can be represented as maximization of a utility function. Utility becomes simply the *name* of the abstract structure over certain choice sets that incorporates the consistency restrictions defined by WARP and Slutsky symmetry. Specifically, a utility function is any mapping from a sequence of choices onto the ordinal line. It is of course the essence of ordinal utility that it does not refer to a *quantity* of anything, such as relative levels of comfort that one might imagine measuring in a person's brain or biochemistry.

Revealed preference is then a purely technical construction: a set of axioms restricting the properties of choices from a set of alternatives that guarantee that the agent's choices can be represented as maximizing a utility function. A modeler can thus be motivated to preserve RPT in order to retain this representational structure, regardless of what latent processes they think might be involved in the generation of choice data.

An axiom of RPT that is of special interest where BR is concerned is completeness. This axiom requires that when we specify an agent by reference to her preferences over identified possible states of some world, we are claiming that there are no two distinguishable states in the identified set for which there is no fact of the matter as to whether the agent prefers one state over the other or is indifferent between them. If we are unsure about some such relation, then our specification of the agent is said to be incomplete with respect to the states in question. If by 'some world' we mean the *entire* world – the set of every consumption bundle the agent could choose if they labored under no information restrictions – then completeness implies denial of BR as I defined it in Section 2. In the classical post-Houthakker version of RPT, the choice set is all possible bundles of all possible consumable goods and services at all possible prices. This set cannot be examined by a finite agent.

However, the point of Samuelson's remark quoted earlier is that the construction of RPT based on this set is a purely conceptual exercise that isn't intended to be of much practical significance. As the uncontroversial BR claim emphasizes, actual choice processes do not operate over infinite sets. But economists were motivated to *make* RPT practically useful by rendering it applicable to finite sets that could represent real data. Thus Afriat (1967), building on Samuelson (1948), showed that one can begin with a finite set of choice data – including, specifically, sequences of real empirical choices by agents at real prices under real budget constraints – and construct a utility function consistent with SARP, but with a purpose different from Houthakker's. The point of Houthakker's theory was to prove existence of a unique utility *maximizing* function. Afriat's aim, by contrast, was to develop an algorithm guaranteed to construct *some* ordinal utility function from choice data. This required a slight relaxation of Houthakker's restrictions. Afriat's Generalized Axiom of Revealed Preference (GARP) conjoins WARP with two exogenous restrictions weaker than Houthakker's, namely monotonicity and convexity of preferences.

In the GARP framework a 'world' is not the entire possible universe of consumption alternatives. It is instead some set of possible states the analyst has chosen to attend to as motivated by some background knowledge, structural theory or pragmatic idealization. An analyst who wishes to investigate a world on which she imposes a stronger version of the REH than Ragot's pragmatic one quoted in Section 2 might use GARP in a way that violates BR. But GARP allows RPT and BR to be reconciled.

There was originally little motivation for identifying any version of RPT, by itself, with the normative honorific of 'rationality,' and Samuelson himself did not.[8] This changed when Savage (1954) built on von Neumann and Morgenstern's (1944) theory of *cardinal* revealed preference to produce the theory of subjective expected utility. (Von Neumann-Morgenstern preferences are still 'revealed' because they do not derive their cardinal values from any theory of latent valuation processes; like Samuelsonian preferences, they summarize observed choice behavior.) Savage certainly *did* think that failure to maximize utility, when information needed to do so is available, is irrational. In my view Savage's methodological shift into a normative philosophy of individual practical action was unfortunate;[9] it is the single most historically important source of the widespread view that economics is an extension of individual decision theory. But Savage's Expected Utility Theory (EUT) does not entail denial of BR because the model of revealed preference it needs is GARP, not SARP. As emphasized by Binmore (2009), Savage insisted that EUT should not be applied to choices in 'large worlds' in which agents lack sufficient information to quantitatively estimate risks. The kind of 'rationality' that is restricted by Savage to applicability in 'small' worlds is a kind of BR; but the bounds in question refer to limits on information flow, not to psychology.

We may contrast Afriat's reconciliation of RPT and BR, motivated simply by recognition of the finitude of data, with the kind of reconciliation proposed by Manzini and Mariotti (2007). They construct an alternative axiom system to WARP (and to GARP) that allows preference structures to be inferred from finite choice sets even when the choices violate restrictions on cycles of the form $a \succ b \succ c \succ a$. One motivation Manzini and Mariotti give for this work is that people not only *do* frequently display such cycles in their choices, but can be expected to *reasonably* display such cycles simply because people express hierarchies of different valuation criteria in the scheduling of their consumption. An economist should be surprised by this motivation only to the extent that she follows Savage in taking the axioms of WARP to represent conceptual analyses of economic rationality – something I have

[8] For a full defense of this claim, see Ross (2014).

[9] An explanation for Savage's rhetoric here is readily available. He was interested in convincing classical statisticians to think rigorously about subjective probabilities, for which purpose appeal to normative intuitions was likely to be helpful or even essential. I thank Glenn Harrison for this observation.

suggested she should avoid. But Manzini and Mariotti offer a second, more speculative, motivation grounded in BR. Citing the frugal heuristics literature, they suggest that people apply sequential hierarchies of decision rules in order to reduce the cognitive load involved in reviewing and selecting from choice sets.

One can imagine situations in which an economist might be led by specific features of a specific market environment she is modeling to use the Manzini and Mariotti axioms in formulating hypotheses for estimating a set of choice data. Even in the case of lottery choice experiments, where outcomes are simply money prizes and so the ordinal structure of utility functions is expected to be simple (all larger monetary quantities preferred by all agents to all smaller quantities) the experimenter might decide to test frugal heuristics hypotheses as a means of investigating heterogeneity in subjects' dispositions to integrate experimental income with their wealth. However, whereas Afriat's achievement motivated a general methodological innovation, this is not a plausible interpretation of Manzini and Mariotti's accomplishment, in light of the *use* to which economists actually put representation theories like GARP in handling data.

RPT is seldom used by economists for predicting out-of-sample choices by individuals. Indeed, prior to the seminal work of McFadden (1974) it seldom featured in applied work at all – as Samuelson anticipated in the remark of his quoted above. McFadden showed how to make RPT relevant to something of much greater interest to economists than individual choices, namely, population-level responses given a distribution of heterogeneous preferences among the population's members. Predictions are derived by modeling preference probabilities as conditional on observed attributes. Where both choice and attribute data are pooled, as is the usual case, one need not suppose that any individual in the population chooses consistently or has a complete preference ordering; one need only suppose that stochastic regularities conforming to GARP associate conditioning variables and choice patterns in the population.

In work of this kind, latent variables are not assumed out of existence. On the contrary, a key area of technical concern and innovation in the econometrics of population-level utility function estimation has been the form of assumed distribution of unobserved attributes. At the time of McFadden's paper, stark limits in computational power forced most applied researchers to impose the assumption of a conditional logit model. However, since the explosion of data-crunching power in the decades since, econometricians have studied a wide range of distributional assumptions. Thus contemporary RPT is a tool for social scale analysis that is sensitive to the play of latent influences on aggregate choice behavior but abstracts away from the dynamics of such influences at the scale of individuals.[10] Heterogeneity is readily extended to the form of the utility function. It is increasingly common for researchers to estimate *mixture models* (Harrison and

[10] To guard against confusion here: the influences in question clearly operate through individual choices, but the *analyses* are insensitive to these channels.

Rutström 2009) designed to identify maximum-likelihood proportions of pooled choices that conform to EUT, RDU – and even, to the extent that an experimental economist is confident that she can induce loss framing by at least some subjects, CPT (Harrison and Swarthout 2013). Individual-scale dynamics of the kind modeled by Manzini and Mariotti might be among the underlying contributing sources of the observed heterogeneity. But to the economist who is professionally interested in market-scale properties rather than individual-scale properties, this is a matter to be left to psychologists. And then, with respect to Manzini and Mariotti's modeling labours, one might pose the question: why try to preserve RPT at the scale of psychological modeling? RPT is for economists, motivated by what economists want to do.

5. Economics and individual rationality

Neo-Samuelsonians emphasize that if there are grounds for associating rationality with conformity of some choices to EUT, the concept of rationality in question is not the philosopher's idea of a normative structure for sound and consistent *thought*. According to neo-Samuelsonians, the most economically important small worlds are *engineered* into being, and maintained, by the institutions, property rights, and coordinated monetary instruments that are necessary conditions for the existence of market structures in which thick personal relationships do not drive out anonymous ones and bad money does not drive out good.[11] Then one can talk about a restricted idea of *market rationality* as restrictions on choices that agents must observe if they are not to place themselves at systematic disadvantage with respect to at least maintaining their wealth in some specified market. To the extent that some such restrictions apply in all or nearly all markets, we might aspire to a *general* theory of market rationality.

Game theory, because it has greatly expanded the range of types of markets that economists know how to model, has weakened confidence in, and the perceived importance of, such a general theory. For example, one could design a game in which an agent *i* should exhibit cyclical preferences over a subset of her preference set in order to deter another agent *j* from taking a game down a path that brought *i* a lower payoff then she would obtain, or stochastically tend to obtain, otherwise.

While general market rationality is of limited and declining interest to economists, the form and extent of market rationality demanded by specific markets and types of markets is of great interest indeed. The scope of variation is not known to be limited. Gode and Sunder (1993) famously simulate a market in which the only market rationality required of agents – or achievable for their simulated 'zero-intelligence traders' – is that they not bid beyond their budget constraints. By contrast, a participant in a derivatives trading market who refuses to use information processing storage technologies that exceed the capacity of her raw brain will almost certainly be driven quickly to extinction in that market.

---

[11] Experimental economists *deliberately* engineer small worlds, usually in the lab but sometimes in the field (Harrison and List 2004).

This last example generalizes. Various authors (Hayek 1945, Satz and Ferejohn 1994, Clark 1997, Martens 2004, Ross 2005) have stressed that people can only make reasonable – let alone optimal – choices in modern markets for goods, services, and investment assets because their institutional environments are replete with what Clark calls 'cognitive scaffolding.' Scaffolded structures embed information computed by prior social processes, and are institutionally maintained and formatted in such a way that they channel choice behavior without requiring choosers to extract the embedded content. Scaffolding guides most human choices in all institutional, networked and normatively structured settings, not just in markets; this is why sociologists or political scientists should no more expect to be able to benefit from a general BR theory than economists. But the forms of scaffolding provided by legally regulated markets are unusually effective in transcending informational restrictions, as Hayek particularly emphasized. To offer just one example, most individual investors could not diversify their savings portfolios by themselves calculating and weighing the fundamental values and risks of the component assets; but they *do* diversify their portfolios by buying types of packages of assets that fund managers are known to be either incentivized or legally obliged to diversify on their behalf. Of course the importance of this aspect of market rationality rests on application of BR to one link in the standard choice-generating process, a biological brain or networked assemblage of biological brains. The expectation that competition and scaffolding will exert powerful effects under most circumstances is the main motivation for Ragot's modest version of the REH as cited in Section 2. But an economist who thinks that in some specific set of circumstances some specific population of agents failed to do what was in their interests (despite conforming to a strategic equilibrium) might typically respect Ragot's rule by citing some consequence of a local manifestation of BR.

Thus the psychology and neuroscience of BR is relevant to economics. But the open possibilities for scaffolding innovations made possible by information-processing technology and institutional engineering impose no known limits on the extent to which market rationality can in principle transcend psychological BR. This is not to claim that there *are* no such limits to be discovered. For example, as noted in Section 2, it is market rational for investors to copy the behaviors of successful investors, under at least some circumstances, even if they do not understand the rationales of the behaviors they replicate. Such herding necessarily results in the loss to the market of some private information, unless no agents ever copy agents who know less than they do. This might give rise to bounds on market rationality that no designable institutions can eliminate. But there are as yet no general limit theorems of this kind, and there may well be no such theorems, applicable to all markets, 'out there' to be discovered.[12] This is entirely compatible with the claim, which should be uncontroversial, that there are no boundless markets.

6. Conclusions

---

[12] For a survey of such general knowledge about herding as had been produced by economists, at least as of a decade ago, see Chamley (2004).

In the foregoing sections I reviewed motivations for the following claims:

(i) The study of the psychology of valuation has arguably revealed some general aspects of BR that are characteristic of choices made by individual people under ecologically unusual circumstances in which tasks are novel, and access to social scaffolding or more experienced agents is strongly constrained. I have not here evaluated the evidence adduced for any such specific hypotheses. The point is rather that this is the interpretive stance from which the classic experimental literature on BR by behavioral economists, at least within the biases and frames tradition, is most charitably assessed.

(ii) The originators of the formal modeling of economic choice did not regard their theory as a model of empirical phenomena. But when RPT *was* adapted for empirical application, it was immediately rendered compatible with local BR hypotheses. Rhetoric about 'grounding' economics in individual decision theory after Savage obscured this point.

(iii) The scope and limits of the rationality of a typical human individual is of secondary interest across most of economics.

(iv) There is no *general* model of BR that applies to the choice phenomena of most interest to economists, in which markets incorporate social and institutional scaffolding; nor should we expect such a theory to be forthcoming. Use of moderate rational expectations assumptions should be understood in this light. Such assumptions are readily relaxed in specific applications in ways that modelers, experimentalists and econometricians are making steadily more sophisticated.

References

Abdellaoui, M., Bleichrodt, H., & Paraschiv, C. (2007). Loss aversion under prospect theory: A paremeter-free measurement. *Management Science* 53: 1659-1674.

Afriat, S. (1967). The construction of utility functions from expenditure data. *International Economic Review* 8: 67-77.

Ainslie, G. (1992). *Picoeconomics*. Cambridge: Cambridge University Press.

Ariely, D. (2008). *Predictably Irrational*. New York: Harper.

Arrow, K. (1951). *Social Choice and Individual Values*. New York: Wiley.

Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C., & Sugden, R. (2010). *Experimental Economics: Rethinking the Rules.* Princeton: Princeton University Press.

Bernheim, B.D., & Rangel, A. (2008). Choice-theoretic foundations for behavioral welfare economics. In A. Caplin and A. Schotter, eds., *The Foundations of Positive and Normative Economics: A Handbook*, pp. 155-192. Oxford: Oxford University Press.

Binmore, K. (2009). *Rational Decisions*. Princeton: Princeton University Press.

Camerer, C., Issacaroff, S., Loewenstein, G., O'Donaghue, T., & Rabin, M. (2003). Regulation for conservatives: behavioral economics and the case for asymmetric paternalism. *University of Pennsylvania Law Review* 151: 1211-1254.

Camerer, C., Loewenstein, G., and Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature* 43: 9-64.

Carlsson, H., & van Damme, E. (1993). Global games and equilibrium selection. *Econometrica* 61: 989-1018.

Chamley, C. (2004). *Rational Herds*. Cambridge: Cambridge University Press.

Clark, A. (1997). *Being There.* Cambridge, MA: MIT Press.

Cox, J., & Harrison, G., eds. (2008). *Risk Aversion in Experiments*. Bingley: Emerald.

Davis, J. (2010). *Individuals and Identity in Economics*. Cambridge: Cambridge University Press.

Gigerenzer, G., Todd, P., & the ABC Research Group (1999). *Simple Heuristics that Make Us Smart.* Oxford: Oxford University Press.

Glimcher, P. (2011). *Foundations of Neuroeconomic Analysis*. Oxford: Oxford University Press.

Gode, D., & Sunder, S. (1993). Allocative efficiency of markets with zero-intelligence traders: market as a partial substitute for individual rationality. *Journal of Political Economy* 101: 119-137.

Gul, F., & Pesendorfer, W. (2001). Temptation and self-control. *Econometrica* 69: 1403-1436.

Hands, W. (2013). Foundations of contemporary revealed preference theory. *Erkenntnis* 78: 1081-1108.

Harrison, G., & List, J. (2004). Field experiments. *Journal of Economic Literature* 42: 1013-1059.

Harrison, G., & Rutström, E. (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental Economics* 12: 133-158.

Harrison, G., & Rutström, E. (2008). Risk aversion in the laboratory. In J. Cox & G. Harrison, eds., *Risk Aversion in Experiments*, pp. 41-196. Bingley, UK: Emerald.

Harrison, G., & Swarthout, T. (2013). Loss frames in the laboratory. In preparation.

Hausman, D. (2011). *Preference, Value, Choice and Welfare*. Cambridge: Cambridge University Press.

Hayek, F. (1945). The use of knowledge in society. *American Economic Review* 35: 519-530.

Houthakker, H. (1950). Revealed preference and the utility function. *Economica* 17: 159-174.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment Under Uncertainty: Heursitics and Biases.* Cambridge: Cambridge University Press.

Kahneman, D. (2011). *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.

Litchtenstein, S., & Slovic, P., eds. (2006). *The Construction of Preference*. Cambridge: Cambridge University Press.

Lucas, R. (1976). Econometric policy evaluation: A critique. In K. Brunner & A. Meltzer, eds., *The Phillips Curve and Labor Markets*, pp. 19-46. London: Elsevier.

Mallard, G. (2011). Modelling cognitively bounded rationality: An evaluative taxonomy. *Journal of Economic Surveys* 26: 674-704.

Manzini, P., & Mariotti, M. (2007). Sequentially rationalizable choice. *American Economic Review* 97: 1824-1839.

Martens, B. (2004). *The Cognitive Mechanics of Economic Development and Institutional Change*. London: Routledge.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka, ed., *Frontiers in Econometrics*, pp. 105-142. New York: Academic Press.

McKelvey, R. & Palfrey, T. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior* 10: 6–38.

McKelvey, R., & Palfrey, T. (1998). Quantal response equilibria for extensive form games. *Experimental Economics* 1: 9–41.

Morris, S., and Shin, H.-S. (2003). Global games: Theory and applications. In M. Dewatripont, L. Hansen and S. Turnovsky, eds., *Advances in Economics and Econometrics, Theory and Applications, Eighth World Congress, Volume 1*, pp. 56–114. Cambridge: Cambridge University Press.

Muth, J. (1961). Rational expectations and the theory of price movements. *Econometrica* 29: 315–335.

Nerlove, M. (1958). *The Dynamics of Supply: Estimation of Farmers' Response to Price* Baltimore: John Hopkins Press.

Nilsson, H., Rieskamp, J., & Wagenmakers, E.-J. (2011). Hierarchical Bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology* 55: 84-93.

O'Donoghue, T., & Rabin, R. (2001). Choice and procrastination. *Quarterly Journal of Economics* 116: 121-160.

Plott, C. (1986). Rational choice in experimental markets. *Journal of Business* 59: S301-S327.

Plott, C., & Zeiler, K. (2005). The willingness to pay - willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. *American Economic Review* 95: 530-545.

Plott, C., & Zeiler, K. (2007). Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory. *American Economic Review* 97: 1449-1466.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization* 3: 323–343.

Ragot, X. (2012). The economics of the laboratory mouse: Where do we go from here? In R. Solow & J.-P. Touffut, eds., *What's Right With Macroeconomics?*, pp. 181-194. Cheltenham: Edward Elgar.

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology, Learning, Memory and Cognition* 34: 1446-1465.

Ross, D. (2005). *Economic Theory and Cognitive Science: Microexplanation*. Cambridge, MA: MIT Press.

Ross, D. (2010). Economic models of procrastination. In C. Andreou & M. White, eds., *The Thief of Time,* pp. 28-50. Oxford: Oxford University Press.

Ross, D. (2012). Neuroeconomics and economic methodology. In J. Davis & W.D. Hands, eds. *The Elgar Companion to Recent Economic Methodology*, pp. 61-93. London: Edward Elgar.

Ross, D. (2014). *Philosophy of Economics*. Basingstoke: Palgrave Macmillan.

Samuelson, P. (1938). A note on the pure theory of consumer's behaviour. *Economica* 5: 61-72.

Samuelson, P. (1947). *Foundations of Economic Analysis.* Enlarged edition (1983). Cambridge, MA: Harvard University Press.

Samuelson, P. (1948). Consumption theory in terms of revealed preference. *Economica* 15: 243-253.

Satz, D., & Ferejohn, J. (1994). Rational choice and social theory. *Journal of Philosophy* 91: 71-87.

Savage, L. (1954). *The Foundations of Statistics.* New York: Wiley.

Simon, H. (1957). *Models of Man: Social and Rational*. New York: Wiley.

Sunstein, C., & Thaler, R. (2003a). Libertarian paternalism. *American Economic Review, Papers and Proceedings* 93: 175-179.

Sunstein, C., & Thaler, R. (2003b). Libertarian paternalism is not an oxymoron. *University of Chicago Law Review* 70: 1159-1202.

Thaler, R. (1992). *The Winner's Curse.* New York: Free Press.

Von Neumann, J, & Morgenstern, O. (1944). *Theory of Games and Economic Behavior.* Princeton: Princeton University Press.

Wakker, P. (2010). *Prospect Theory for Risk and Uncertainty*. Cambridge: Cambridge University Press.

Yaari, M. (1987). The dual theory of choice under risk. *Econometrica* 55: 95-115.