Digit Ratios and Social Preferences: A Comment on Buser (2012)*

Pablo Brañas-Garza† Jaromír Kovářík‡

December 20, 2013

Abstract

Buser (2012) reports an association between the second-to-fourth digit ratio, a biomarker of the exposure to prenatal sex hormones, and behavior in several classic experimental games designed to elicit prosocial attitudes. His subjects self-report whether they have shorter, equal, or larger ring than index finger. We argue that this elicitation method is inappropriate. It generates a poor proxy for the digit ratio as it suffers from measurement errors. As a result, using this variable in the regression analysis may lead to inconsistent estimates.

Keywords: Digit ratio, measurement errors, endogeneity, social preferences, non-monotonicity, altruism.

1 Introduction

An increasing number of studies in many behavioral disciplines analyze the relation between the second-to-fourth digit ratio (2D:4D, hereafter) and behavioral, cognitive, and health-related human characteristics (see Voracek and Loibl (2009) for a review). The 2D:4D (the ratio between the lengths of the index and ring fingers) is a biomarker of exposure to fetal sex hormones, which have permanent organizing effects on the human brain (Lutchmaya *et al.*, 2004; Zheng and Cohn, 2011).

*We are indebted to Burkhard Schipper, Robert Slonin and Santiago Sanchez-Pagés for sharing their data. We also thank to Coralio Ballester, Antonio Cabrales, Antonio Espín, María Paz Espinosa, Nikolaos Georgantzís, Nagore Iriberri, Praveen Kujal, Debrah Meloso and Martin Voracek for extremely helpful comments. Brañas-Garza acknowledges financial support from the Spanish Ministry of Science and Innovation (ECO2010-17049), Junta de Andalucía-Excellence (P07-SEJ-02547), and Fundación Ramón Areces (R+D 2011). Jaromír Kovářík acknowledges support from the Basque Government (IT-783-13) and the Spanish *Ministerio de Economía y Competividad* (ECO 2012-31626, ECO 2012-35820).

†Middlesex University London, Business School, Hendon Campus, The Burroughs, London NW4 4BT, UK (branasgarza@gmail.com).

‡Dpto. Fundamentos Análisis Económico I & BRiDGE, University of the Basque Country, Av. Lehendakari Aguirre 83, 48015 Bilbao, Spain (jaromir.kovarik@ehu.es).

In the last decade, many papers have studied the association between 2D:4D and economic preferences and behaviors, such as risk aversion (Sapienza *et al.*, 2009, Garbarino *et al.*, 2011; Brañas-Garza and Rustichini, 2011), performance of financial traders (Coates *et al.*, 2009), and bidding behavior in auctions (Pearson and Schipper, 2012) among others.[1] Since both null (e.g. Apicella *et al.,* 2008; Drichoutis and Nayga, 2012; Schipper, 2012) and non-monotonic associations (e.g. Sanchez-Pages and Turiegano, 2010; Nye *et al.*, 2012; Brañas-Garza *et al.*, 2013) between 2D:4D and some economically relevant outcomes are not uncommon, precise measurements of 2D:4D are crucial for the detection of the correct associations.[2]

In volume **76**, *Games and Economic Behavior* published a paper relating 2D:4D and behavior in Ultimatum, Trust, Public Goods, and Dictator Games (Buser, 2012). A critical element of this paper is that subjects' hands are not measured using standard techniques. Rather, subjects self-reported whether they had the ring finger larger/equal/shorter than the index finger. The author claims that *"While this measure is less precise than, for example, hand scans, it has a number of advantages. It can be easily included in any questionnaire and is less intrusive on the privacy of subjects"* (p. 459) and validates this elicitation method using a sample of 78 different subjects (45 women), reporting highly significant correlations between the scanned 2D:4D and such self-reports.

In this comment, we would like to illustrate that this method is highly inappropriate and, contrary to Buser's recommendation and validations, should not be used in future research. Compared to the literature, the elicited distribution of 2D:4D in Buser (2012) is misrepresented (even if we accept his coarse categorization) and it may lead to endogeneity problems in the regression analysis. As a result, the methodology may suggest incorrect associations or association that do not exist between 2D:4D and the variable of interest.

2 Digit ratio elicitation and distribution

First of all, Buser (2012, p.459) concludes that his self-reported coarse measure is a good proxy for 2D:4D. However, 2D:4D is already a proxy itself, a proxy for the exposure to fetal hormones. Hence, even if we accept Buser's (2012) validation, his measure represents *a proxy variable of a proxy variable* and we believe that any use of such proxies of proxies deserves a more delicate analysis.

As for the self-assessment of digit lengths, the scale and standard deviations of 2D:4D are typically too small to be easily captured with the naked eye. In fact, measurement methodology studies (Kemper and Schwerdtfeger, 2009; Voracek *et al.*, 2007) even discourage low-resolution measurement methods due to their low reliability. In addition, self-reported measures are inherently unreliable. For instance, it is possible that subjects aware of the relation between 2D:4D and

[1]See Millet (2011) for a survey of associations between 2D:4D and economic decision-making.

[2]Voracek *et al.* (2007) discuss the measurement issues more systematically.

sex differences may over or under-report their 2D:4D to signal (consciously or unconsciously) their femininity or masculinity.

Even if we accept the three-category approximation of 2D:4D in Buser (2012), his data contain an excessive number of people reporting 2D:4D higher than or equal to one. In particular, 38.9% of Buser's (37.6% of the European) subjects report larger index than ring fingers (i.e. a ratio larger than one) on the right hand; meanwhile 14.1% (15.4%) report a ratio equal to one (Buser, 2012, p. 461).[3]

We compare these figures with four independent data sets. The data come from three recent studies unrelated to the authors of this comment: Sanchez-Pagés and Turiegano (2010; SPT, henceforth) accounting for a sample of 147 Caucasian males in Europe, Garbarino *al.* (2011; GSS) reporting 2D:4D of 151 Caucasians (43.04% women) in Ohio, and Pearson and Schipper (2012; PS) analyzing 400 Caucasian and non-Caucasian subjects (46.75% females) in California. We add one larger data set from our research (Bosch-Domenèch *et al.*, 2013; BDBE), accounting for 653 Caucasian Spanish participants (58.04% women).[4] In all cases, subjects' digit lengths were elicited using standard techniques in the literature (see Voracek et al. (2007) for a review of measurement methods and their precision): participants' hands were scanned, their fingers were independently measured several times from the basal crease to the tip of the finger, and the average of all measurements was used in the analysis.[5] In these studies, 10.9% (SPT), 8.6% (GSS), 6.3% (PS), and 16.8% (BDBE) of participants have 2D:4D larger than one. Only 0.3% of participants exhibit $2D : 4D = 1$ in BDBE, whereas nobody does in the other three distributions.

To compare these data sets with Buser (2012), we created a variable that is equal to 1 (0 and -1) if 2D:4D is larger than (equal to and lower than) one. Figure 2 contrasts the European subjects in Buser (2012) with the other distributions. Observe that the fractions of individuals reporting both $2D : 4D = 1$ and $2D : 4D > 1$ are systematically larger in Buser (2012) than in the other cases, whereas the latter cases are mutually comparable. Two-sided Wilcoxon rank sum tests reject the equality of Buser's (2012) data with these four alternative distributions at any significance level in any case ($z > 5$; $p = 0$ in all cases).[6] Hence, people either over-report their 2D:4D's or report a different feature of their hands in Buser (2012).

We provide two additional tests. First, we allow for mistakes in the BDBE

[3]There is a total of 252 (221 European) subjects in Buser (2012), being 62.3% female.

[4]Since 2D:4D is subject to important ethnic variation (e.g. Manning *et al.*, 1999, 2004), we compare Caucasians with Buser's (2012) European subjects below. One exception is PS, who use a mixture of Caucasian and non-Caucasian subjects. We opted to use the mixed population in PS to illustrate that the documented ethnic differences cannot account for such a big difference between Buser (2012) and the other distributions. Our conclusions are unaffected if we include non-European subjects from Buser (2012) into the analysis.

[5]More details about the samples and measurement methods can be found in the corresponding articles. All these studies use right hands of subjects, except Garbarino *et al.* (2011) who use the average of the left- and right-hand 2D:4D. This is the measure we use here.

[6]In contrast, we cannot reject the pairwise equality of the five alternative data sets in all cases ($p > 0.1$), with two exceptions: BDBE vs. PS and BDBE vs. GSS ($p = 0$ and 0.01, respectively).

data and set 2D:4D as higher (lower) than one if it is higher (lower) than 1.01 (0.99). This shifts the shares of the $2D : 4D < 1$ and $2D : 4D = 1$ participants towards Buser (2012) but at the expense of the $2D : 4D > 1$ subjects and the distribution is still statistically different from Buser (2012; $z = 8.381$; $p = 0$). Second, note that there are more women in Buser (2012) than in the other data sets. To match the gender composition of Buser (2012), we create 100 random samples containing 62.3% using the BDBE data and again created the corresponding categorical variable, included in Figure 1. The 95% confidence intervals for e.g. $2D : 4D < 1$ are $(0.8243; 0.8255)$, way above Buser's (2012) 47% of European subjects.[7] Hence, neither the above reporting errors nor gender can account for the observed differences.

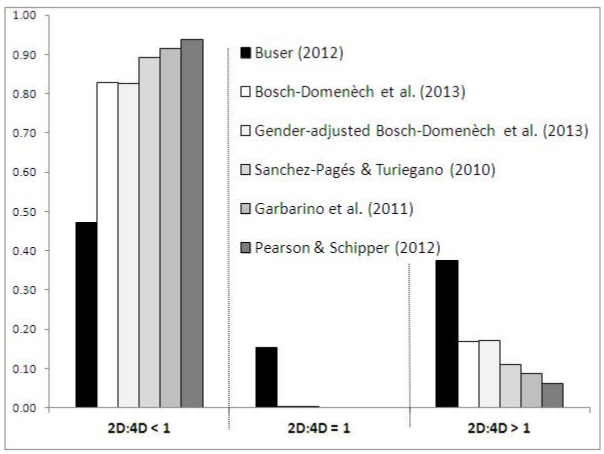

Figure 1: Distributions of right-hand 2D:4D of Europeans in Buser (2012) and the discretisized distributions from Bosch-Domenèch *et al.* (2013), the gender-adjusted Bosch-Domenèch *et al.* (2013), Sanchez-Pagés and Turiegano (2010), Garbarino *et al.* (2011), and Pearson and Schipper (2012).

Voracek *et al.* (2008) analyze the relationships between 2D:4D and *finger tip prominence*, measuring the prominence of the ring with respect to the index finger using three categories (more prominent, less prominent, and an intermediate case). This measure abstracts from the position of the basal crease of the fingers and thus represents an inappropriate measure of 2D:4D. Indeed, Voracek *et al.* (2008) conclude that, even if they are correlated, the *finger tip prominence is only a weak proxy for actual 2D:4D*, since it only accounts for 25% of the interindividual variation in 2D:4D. Since Buser's (2012) self-reported measure resembles finger tip prominence in many respects, it seems that Buser's (2012) subjects report this variable.

[7]Since all the 95% confidence intervals of the categorical variable corresponding to the simulated gender-adjusted distributions are very small, we do not include them in Figure 1.

3 Discussion

Exploring biological roots of economic preferences is an important step towards the understanding of economic behavior, and growing evidence suggests that exposure to fetal hormones (measured by 2D:4D) may represent one such biological factor (Manning et al., 2002; Voracek and Loibl, 2009). Measuring 2D:4D through casual inspection, as suggested by Buser (2012), might be easier to administer in a lab. However, as illustrated in this comment this procedure may lead to serious misrepresentations of the true distributions of this variable.

The most severe consequence of this misrepresentation is the inconsistency of estimates when the errors-in-variables problem is not treated properly (see e.g. Wooldridge, 2002, Chapter 4).[8] If the relation between the dependent and independent variables is linear and certain assumptions concerning the measurement error hold, the "only" problem may be the attenuation bias. However, recent evidence suggests non-linear associations between 2D:4D and behavior in the Prisoner's Dilemma and the Dictator Game (Sanchez-Pagés and Turiegano, 2010; Brañas-Garza *et al.*, 2013). If the true association is indeed non-linear or if the assumptions leading to the attenuation bias do not hold, the biases in the estimates can go to any direction. As a result, it cannot be concluded from Buser (2012) that prenatal hormones affect subjects' behavior in his experiment.

Hence, we cannot recommend such elicitation methodology of 2D:4D to other scholars. All the issues discussed in this comment can be easily overcome by scanning of subjects hands using high-resolution scanners.

References

[1] Apicella, C. L., Dreber, A., Campbell, B., Graye, P. B., Hoffman, M., Littleget, A. (2008) Testosterone and financial risk preferences, *Evolution and Human Behavior* 29, 384–390.

[2] Bosch-Domenèch, A., Brañas-Garza P., and A. Espín (2013), Fetal testosterone (2D:4D) as predictor of cognitive reflection, mimeo, Universitat Pompeu Fabra.

[3] Brañas-Garza P, Kovářík J, Neyse L (2013) Second-to-Fourth Digit Ratio Has a Non-Monotonic Impact on Altruism, *PLoS ONE* 8, e60419.

[4] Brañas-Garza, P., Rustichini, A. (2011) Organizing effects of Testosterone and Economic Behavior: Not Just Risk Taking, *PLoS ONE* 6, e29842.

[5] Buser T. (2012) Digit ratios, the menstrual cycle and social preferences, *Games and Economic Behavior* 76, 457-470.

[6] Coates, J., Gurnell, M., Rustichini, A. (2009). Second-to-fourth digit ratio predicts success among high-frequency financial traders, *Proceedings of the National Academy of Sciences of the USA* 106, 623-628.

[8] Consequently, increasing the sample size does not solve the issue.

[7] Drichoutis, A., Nayga, R. (2012) Do risk and time preferences have biological roots?, mimeo, University of Akransas.

[8] Garbarino, E., Slonim R., Sydnor J. (2011) Digit ratios (2D:4D) as predictors of risky decision making for both sexes, *Journal of Risk and Uncertainty* 42, 1–26.

[9] Kemper CJ, Schwerdtfeger A (2009) Comparing indirect methods of digit ratio (2D:4D) measurement, *American Journal of Human Biology* 21, 188–191.

[10] Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., Manning, J.T. (2004) 2nd to 4th digit ratios, fetal testosterone and estradiol, *Early Human Development* 77, 23-28.

[11] Manning JT, Stewart A, Bundred PE, Trivers RL. (2004) Sex and ethnic differences in 2nd to 4th digit ratio of children, *Early Human Development* 80, 161-168.

[12] Manning JT, Trivers RL, Singh D, Thornhill R (1999) The mystery of female beauty, *Nature* 399, 214-215.

[13] Millet K (2011) An interactionist perspective on the relation between 2D:4D and behavior: An overview of (moderated) relationships between 2D:4D and economic decision making, *Personality and Individual Differences* 51, 397-401.

[14] Nye JVC, Androuschak G, Desierto D, Jones G, Yudkevich M (2012) 2D:4D Asymmetry and Gender Differences in Academic Performance, *PloS ONE* 7, e46319.

[15] Pearson, M., Schipper, B. C. (2012) The visible hand: finger ratio (2D:4D) and competitive bidding, *Experimental Economics* 15, 510-529.

[16] Sanchez-Pages S, Turiegano E (2010) Testosterone, facial symmetry and cooperation in the prisoners' dilemma, *Physiological Behavior* 99, 355–361.

[17] Sapienza, P., Zingales, L. and Maestripieri, D. (2009) Gender differences in financial risk aversion and career choices are affected by testosterone, *Proceedings of the National Academy of Sciences of the USA* 10.1073/pnas.0907352106.

[18] Schipper, B. (2012) Sex hormones and choice under risk, mimeo, University of California, Davis.

[19] Voracek, M, Dressler SG, Loibl LM (2008) The contributions of Hans-Dieter Rösler: Pioneer of digit ratio (2D:4D) research, *Psychological Reports* 103, 899-916.

[20] Voracek M, Loibl LM (2009) Scientometric analysis and bibliography of digit ratio (2D:4D) research, 1998-2008, *Psychological Reports* 104, 922-956.

[21] Voracek M, Manning JT, Dressler SG (2007) Repeatability and interobserver error of digit ratio (2D:4D) measurements made by experts, *American Journal of Human Biology* 19, 142-146.

[22] Wooldridge, J. M. (2012). *Introductory econometrics: a modern approach.* Cengage Learning.

[23] Zheng Z, Cohn MJ (2011) Development basis of sexually dimorphic digit ratios, *Proceedings of the National Academy of Sciences of the USA* 108, 16289–16294.