Working Paper in Economics No. 750

The digit ratio (2D:4D) and economic preferences: no robust associations in a sample of 330 women

Elle Parslow, Eva Ranehill, Niklas Zethraeus, Liselott Blomberg, Bo von Schoultz, Angelica Lindén Hirschberg, Magnus Johannesson, and Anna Dreber

Department of Economics, February 2019

ISSN 1403-2473 (Print)
ISSN 1403-2465 (Online)

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

The digit ratio (2D:4D) and economic preferences: no robust associations in a sample of 330 women*

Elle Parslow[1], Eva Ranehill[2], Niklas Zethraeus[3], Liselott Blomberg[4], Bo von Schoultz[4], Angelica Lindén Hirschberg[4], Magnus Johannesson[1], and Anna Dreber[†1]

[1] *Department of Economics, Stockholm School of Economics*
[2] *Department of Economics, University of Gothenburg*
[3] *Department of Learning, Information, Management and Ethics, Karolinska Institutet*
[4] *Karolinska University Hospital, Karolinska Institutet*

January 2019

Abstract

Many studies report on the association between 2D:4D, a putative marker for prenatal testosterone exposure, and economic preferences. However, most of these studies have limited sample sizes and test multiple hypotheses (without preregistration). In this study we mainly replicate the common specifications found in the literature for the association between the 2D:4D ratio and risk taking, the willingness to compete, and dictator game giving separately. In a sample of 330 women we find no robust associations between any of these economic preferences and 2D:4D. We find no evidence of an effect for sixteen of the eighteen total regressions we run. The two regression specifications which are significant have not previously been reported and the associations are not in the expected direction, and therefore they are unlikely to represent a real effect.

Keywords: 2D:4D; economic preferences; experiment; testosterone
JEL codes: C91, D03

*We thank David Bilén for research assistance, and Levent Neyse, Pablo Brañas-Garza and Thomas Buser for helpful comments. This work was supported by research grants from the Jan Wallander and Tom Hedelius Foundation (Grants P2010–0133:1, P2012–0002:1, P2013–0156:1, P2017-0143:1, and H2015-0408:1), the Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellows grant to A. Dreber), the Swedish Council for Working Life and Social Research (Grant 2006–1623), the Swiss National Science Foundation (Grant 100010–149451), the Swedish Research Council (Grant 20324), Karolinska Institutet, and the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet (Grant 20130313).

†To whom correspondence and requests for materials should be addressed: Anna Dreber, Department of Economics, Stockholm School of Economics, P.O Box 6501, SE-11383 Stockholm, Sweden. Phone: +46 8 7369446. Email: anna.dreber@hhs.se

1 Introduction

Testosterone has been hypothesised to be associated with a wide range of economic decision making. One aspect of this hypothesis is the theory that prenatal testosterone exposure impacts brain development and therefore can explain some of the heterogeneity in behaviour between individuals. A putative proxy for the level of prenatal testosterone exposure is the ratio of the length of the 2nd digit to the length of the 4th digit (2D:4D) on each hand, as suggested by Manning et al. (1998). Subsequently, many studies have reported associations between 2D:4D and a variety of traits, such as sexual orientation, spatial ability and personality traits, although the results are often conflicting (and with some possibility of publication bias, see e.g. Puts et al. (2008), Voracek and Loibl (2009), Grimbos et al. (2010), Voracek et al. (2011), but see Hönekopp and Schuster (2010) and Hönekopp and Watson (2011), who do not find evidence for publication bias). Furthermore, a sizeable literature uses 2D:4D to explore the effect of prenatal testosterone exposure on economic decisions, also with mixed results.

This paper aims to test hypotheses in previous papers in relation to the association between 2D:4D and risk taking, dictator game giving, and the willingness to compete. These preferences are relevant for explaining variation in many economic outcomes. We use a sample of 330 women – which is large given most sample sizes that have previously been reported – in an experiment to measure 2D:4D and economic preferences. For each of our three outcome measures, we run separate regressions for the left hand, the right hand and the average of both hands, with linear and quadratic specifications, in a replication of what is mainly used in the literature. In the nine linear regression models tested we find no evidence of an effect ($p > 0.05$). In the nine regressions with a quadratic term we find significant evidence ($p < 0.005$) in two regressions. However, the direction of the effect in these two regressions is not in line with previous hypotheses (and has not been reported before in the literature), and is therefore not likely to be a real effect. Overall, we conclude that we find no evidence for an association between 2D:4D and the economic preferences measured in this sample of women.

Whilst the 2D:4D measure has been used in many studies, the link between prenatal testosterone and 2D:4D is not strongly established (McIntyre (2006)). Evidence from humans and other mammals show organisational effects on the brain of hormones (such as testosterone) during foetal development (Phoenix et al. (1959), Arnold (2009), Lombardo et al. (2012)), which may affect behaviour and preferences later in life and is thus investigated in many 2D:4D studies. The 2D:4D ratio is likely set early in utero stimulated by androgens on the androgen receptor ((Lutchmaya et al., 2004), Zheng and Cohn (2011), Galis et al. (2010)), as there are sex differences in foetal sex steroid hormone levels as measured in amniotic fluid (Lutchmaya et al. (2004), van de Beek et al. (2004), Ventura et al. (2013)) while about six months after birth there are no longer any sex differences in sex hormone levels until puberty (Abreu and Kaiser, 2016). In adult humans, 2D:4D has been shown to be lower in males than females, although the majority of research is conducted on Western populations (c.f. Apicella et al. (2016), who finds no sex difference in 2D:4D in a remote, hunter-gatherer population). The oft-cited study by Lutchmaya et al. (2004) which investigates the link between 2D:4D and prenatal testosterone exposure, finds a significant negative correlation in a sample of 29 children between the testosterone-to-estradiol ratio in amniotic fluid and right hand 2D:4D only, even after controlling for gender (the left hand is reported insignificant). An additional method of investigation is to compare same sex and opposite sex twins, based on the theory of sex-hormone transfer in utero (Miller (1994)). van Anders et al. (2006) find that females with a male rather than female co-twin have lower left hand 2D:4D, which the authors argue is due to hormone transfer from male to female foetuses, however they find no significant results for the right hand. Whilst Voracek and Dressler (2007) in a similar study report a significant result for mean 2D:4D, among studies with much larger sample sizes there is a failure to find significant differences (Hiraishi et al. (2012), Cohen-Bendahan (2005), Medland et al. (2008)).[1] Other methods of establishing a link between 2D:4D and androgen exposure both post- and peri-natally include using congenital adrenal hyperplasia (CAH) and the CAG repeat polymorphism (McIntyre (2006), Brown et al. (2002)), and here also there is a mix of positive and null results.

Even though the link between 2D:4D and prenatal testosterone is not well established, there are many papers investigating the association of 2D:4D with economic decision making. Whilst 2D:4D is an easy-to-measure way to proxy for prenatal testosterone exposure, many of these papers use multiple tests and have relatively small sample sizes. As far as we are aware, none of the previous studies pre-register their analyses. There are often multiple hypotheses involving different ways of measuring the explanatory variable (left hand, right hand, average of both hands or even squared 2D:4D), as well as which controls to include (such as gender, age or sexual orientation) and which subsamples to analyse (such as ethnicity and gender), giving rise to many 'forking paths' (Gelman and Loken (2013)) and researcher degrees of freedom (Simmons et al. (2011)). As discussed in Simmons et al. (2011), researchers have many options available in choosing among outcome variables, controls and subsample selection, creating ambiguity in the research process and potentially generating higher rates of false positives than 5%, even if researchers do not intend to do so. In our review of the literature in the following subsections, we consider significant results to be cases where the p-value is less than 0.05 and report anything above that threshold as insignificant, as is typically used. We present tables to summarise the results of studies that use comparable measures of economic preferences to our experiments.[2] However in our own results in this paper, we instead consider a p-value less than 0.05 to indicate suggestive evidence, whilst statistical significance requires a p-value less than 0.005, following Benjamin et al. (2018).

Benjamin et al. (2018) suggest a change in the p-value defining statistically significant new discoveries from 0.05 to 0.005, to improve the reproducibility of scientific studies (in terms of reducing rates of false positives). The authors propose that where p-values are below 0.05 but above 0.005, this should be interpreted as suggestive evidence. Whilst our study aims to be a replication of past studies, the results of past studies are mixed and therefore we think it is appropriate to use the more conservative 0.005 threshold for statistical significance. An additional motivation for a more conservative threshold than 0.05 is that we, following the existing literature, run several tests for each outcome measure.

1.1 Dictator game giving

Several papers have looked at the relationship between 2D:4D and giving in the dictator game.[3] The dictator game removes any repercussions of failure to reciprocate (unlike the ultimatum game), and in all the below studies the participants were told that the recipient in the game is another participant whose identity is unknown.[4] The hypothesised relationship between 2D:4D and dictator game giving is positive, with higher exposure to testosterone (low 2D:4D) being associated with lower levels of

[1] Studies finding significant results have sample sizes of 24 and 28 (van Anders et al. (2006) and Voracek and Dressler (2007) respectively) whereas studies finding no significant differences report sample sizes of 55, 55 and 449 (Cohen-Bendahan (2005), Hiraishi et al. (2012) and Medland et al. (2008), respectively).

[2] The results we include in the table for mixed gender samples are specifications which include a gender control only, unless noted otherwise.

[3] Some studies have looked at other games, such as public goods games, and interpret behaviour as altruistic or pro-social, such as Millet and Dewitte (2006)

[4] However Millet and Dewitte (2009) use a hypothetical dictator game.

Study	Sample	Men						Women						Both sexes					
		L	L L^2	R	R R^2	M	M M^2	L	L L^2	R	R R^2	M	M M^2	L	L L^2	R	R R^2	M	M M^2
Millet and Dewitte (2009), Study 1 neutral prime	119ˆ															S-			
Millet and Dewitte (2009), Study 1 aggressive prime	119ˆ															NS			
Millet and Dewitte (2009), Study 2 neutral prime	90ˆ															NS			
Millet and Dewitte (2009), Study 2 aggressive prime	90ˆ															S+			
Buser (2012)	221													NS		NS**		NS**	
Brañas-Garza et al. (2013), 2010 study	()		S+, S- (87)		S+, S- (88)				NS, NS (61)		S+, S- (61)			NS* (170)	S+, S- (170)	NS* (171)	S+, S- (171)		
Brañas-Garza et al. (2013), 2011 study	()		NS, NS (68)		NS, NS (69)				NS, NS (53)		S+, S- (53)			NS* (126)	NS, NS (126)	S+* (127)	S+, S- (127)		
Galizzi and Nieboer (2015), all	602													NS	NS, NS	NS	NS, NS		
Galizzi and Nieboer (2015), Caucasian	201													NS	NS	NS	S+, S-		
Galizzi and Nieboer (2015), Chinese	221													NS	NS	NS	NS		
Galizzi and Nieboer (2015), South Asian	81													NS	NS	NS	NS		

L: left hand. R: right hand. M: mean of left and right hands.
NS: not significant. S+: significant positive relationship. S-: significant negative relationship.
*Does not control for gender; **significant positive result for binary variable where 1 indicates 4D longer and 0 indicates all other scenarios (i.e. same as 2D or shorter), authors note results do not change for genders separately, they also control for age, nationality and experience in previous games; ˆsample size is the total sample for that study, the authors do not state the sample split for neutral or aggressive prime groups

Table 1: Dictator game giving studies

dictator game giving. The results from studies using the dictator game are summarised in Table 1, showing that non-significant findings are common. When significant, regressions using squared 2D:4D measures find an inverse U-shaped relationship between 2D:4D and dictator game giving (low dictator game giving is associated with both low and high testosterone).

In a study looking at several social preference games, Buser (2012) runs a dictator game experiment and measures a binary indicator for choosing the selfish allocation compared to the altruistic allocation. This study uses a self-reported measure of 2D:4D, which has been criticised due to measurement error (see, e.g. Brañas-Garza and Kovarik (2013)). When splitting the sample into three groups with a dummy for low 2D:4D and a dummy for high 2D:4D (the reference group is therefore ring finger equal to index finger), the two dummy variables are not jointly significant, for either hand. However there is a positive correlation between selfishness and the right hand ring finger being longer than the index finger (low 2D:4D), relative to the rest of the sample. The study finds the same positive correlation when both left and right hand ring finger is longer, but an insignificant result for just the left hand. Thus this study only finds significant results when comparing low 2D:4D subjects to the rest of the sample (composed of intermediate and high 2D:4D).

While still in the context of a dictator game, Millet and Dewitte (2009) use a modified experimental strategy by priming mood (neutral or aggressive) in subjects and then measuring hypothetical dictator game giving, in a mixed gender sample. In the variation with a music video prime, they find a significant negative association between right-hand 2D:4D and dictator game giving under the neutral condition and no relation under the aggressive condition (sample sizes for each of the conditions were unreported). In the second variation, using a language task with aggressive or neutral words, the result reverses with a positive correlation under the aggression condition and no result under the neutral condition.

Brañas-Garza et al. (2013) introduce the use of squared 2D:4D measures in this context and find conflicting results. Dictator game giving is significantly positively associated with male, female and combined gender samples of right-hand 2D:4D, and in left hand 2D:4D in males and combined genders only. It is negatively correlated with male and both genders left and right hand squared 2D:4D and female right hand squared 2D:4D. They run the game again one year later with the same subjects but slightly smaller sample, and find less significant results, as all effects have disappeared for the male only sample, and the left hand and left hand squared effect is insignificant for the combined gender sample. Right hand 2D:4D becomes significant for the combined gender sample, showing a positive relationship with dictator game giving (although gender is not reported as a control in this regression).

Galizzi and Nieboer (2015) look at the digit ratio and dictator game giving in the context of different ethnic groups. They do not find a significant correlation between dictator game giving and digit ratio in the full sample. Splitting the sample by ethnicity (Caucasian, Chinese and South Asian) and controlling for gender, they run OLS regressions on left and right hand 2D:4D separately, and then add the respective squared values of each in further regressions, resulting in 16 regression specifications. Of these, only the Caucasian subsample finds a significant relationship, with a positive coefficient for right hand 2D:4D and a negative coefficient for right hand 2D:4D squared (the regression with right hand 2D:4D alone is not significant).

1.2 Risk taking

While several review papers find that women are on average more risk averse than men, (see, e.g. Eckel and Grossman (2008), Croson and Gneezy (2009), Charness and Gneezy (2012)), there is also evidence from a meta-analysis by Nelson (2015) suggesting that the difference (in terms of effect size) is not very large. Nevertheless, there is a substantial literature looking into a biological explanation for this gender difference through prenatal testosterone exposure and the 2D:4D ratio. As far as we are aware, only one study finds an association between 2D:4D and risk tasking in men and not in women (Stenstrom et al., 2011). The hypothesis is that risk taking is negatively related to 2D:4D — higher testosterone exposure is associated with higher risk taking (and lower risk aversion). We limit our analysis of the previous literature to the areas of financial or general risk taking. There are numerous ways to measure risk-taking in experimental tasks, as well as the digit ratio (such as by scanner, or calliper etc.), which can add measurement error.

In an early paper investigating risk preferences and 2D:4D, Dreber and Hoffman (2007) measure risk-taking using a one-off risky investment decision. They report correlations between 2D:4D and risk for a Swedish sample, finding left hand 2D:4D is negatively correlated with risk taking after controlling for gender and ethnicity, but find no relationship for the right hand. However in a second American sample, they find no evidence for either hand. Also using a one-off risky investment decision, Apicella et al. (2008) find no significant correlation between either left or right hand 2D:4D and risk preferences in a smaller, all male, sample. Subsequent literature has explored a variety of risk taking measures, as well as heterogeneous effects of ethnicity or gender using subsamples, and different specifications of the explanatory variables and controls.

Roughly half of the studies report all insignificant results, across a variety of 2D:4D specifications (left hand, right hand or average of both hands) and with some using subsamples (split by gender or ethnicity). Sapienza et al. (2009) find no significant relationship between risk preferences and average 2D:4D in a mixed gender sample of 181. Running separate regressions for men and women, or for right and left-hand, also leads to insignificant results. Aycinena et al. (2014) look at an ethnically homogeneous sample from Guatamala and find no significant association for either hand, and neither using specifications with squared 2D:4D. Splitting into subsamples by gender also yields insignificant results. Schipper (2014) measures only right hand 2D:4D, and risk preferences are measured in both the gain and loss domains (to control for framing effects) but there are no significant results for men or women, even when using ethnically homogeneous subsamples (although these sample sizes are small). Chicaiza-Becerra and Garcia-Molina (2017) use a sample from Colombia and follow the approach of Garbarino et al. (2011), but find no significant correlations for either right or left hand 2D:4D and risk taking for both the full sample and a more ethnically homogeneous subsample. Lima de Miranda et al. (2018) find no significant results for either left or right hand 2D:4D, as well as left and right hand 2D:4D squared. Finally, Alonso et al. (2018) assemble data from five experimental projects and the 2D:4D measure they use is a dummy variable which takes the value 1 if the subject has a 2D:4D measure above the gender-specific median value, and 0 otherwise. They look at the relationship between high 2D:4D and risky choices only for those subjects who made consistent decisions in the risk preference task, and find no significant relationship for either right or left hand 2D:4D.

Other studies test multiple hypotheses and report mixed results, showing some significant relationships with the sign in the hypothesised direction along with some insignificant findings. Sytsma (2014) uses the same methodology to measure risk preferences as in Garbarino et al. (2011), using a sample from Bangladesh. Two lotteries are employed: one with gains framing, the other with loss framing. Then as an additional measure, the average of the individual's choices over the 2 lotteries is used, and the digit ratio is tested for the left hand, right hand and average of both hands. The author finds significance in the same direction (negative association between 2D:4D and risk taking) for 6 out of the 27 specifications tested. In a study of 211 Israeli students, Barel (2017) looks at risk-taking moderated by optimism using self-reported risk measures, and finds a negative correlation between right hand 2D:4D and general risk taking, but no association for the left hand or for financial risk taking (2D:4D is standardised within gender). Also using self-reporting, Stenstrom et al. (2011) use survey based measures of risk taking and measure right hand 2D:4D. They find a significant negative correlation between financial risk taking and 2D:4D, and overall risk taking and 2D:4D, for the male subsample only.[5] They also look at a Caucasian subsample and find significance only for Caucasian males between overall risk taking and right hand 2D:4D. All results for females are insignificant.

Some studies limit to only one specification (such as one measure of 2D:4D and one risk taking measure using the full sample, although controls may differ) relating to 2D:4D and risk taking, although results are mixed. Garbarino et al. (2011) use normalised mean 2D:4D in a mixed-gender sample, controlling for both gain or loss framing and for other risk measures, and report one specification, finding a significant negative relationship between 2D:4D and a risky choice task.[6] In a field experiment, Coates and Page (2009) look at a small sample of male traders. Using the standard deviation of their profit and loss over a 20 month period as the risk measure, they find that risk taken correlates negatively with right hand 2D:4D. In contrast, Drichoutis and Nayga (2015) use right hand and right hand squared 2D:4D and run a single regression specification. Their result finds no significant relationship with risk taking, in a mixed-gender sample from Greece.

Whilst the hypothesised relationship between risk taking and 2D:4D is negative, one study finds a positive correlation. Brañas-Garza and Rustichini (2011) use two lottery choice tasks to create two measures of risk aversion for 188 caucasian subjects. Risk taking is positively correlated with 2D:4D for the female subsample of 116 participants, but insignificant for the male subsample. Their combined risk aversion (CRA) measure is insignificant for both genders, in a regression of CRA on right hand 2D:4D and a constant. In a simple correlation, CRA is positively correlated with 2D:4D for males only, not for females or the combined sample.

Two recent studies have the largest sample sizes across studies relating 2D:4D to risk preferences. Bönte et al. (2016) use a survey questionnaire on willingness to take risks in general, and in investment, instead of an experimentally elicited risk preference. They find that 2D:4D is negatively related to general risk taking but not to investment risk taking, for the right hand.[7] They find left hand 2D:4D to be uncorrelated with risk so do not report the full regression results for the left hand. In a larger study, Brañas-Garza et al. (2017) use both self-reported risk attitude and lottery choices to measure risk preferences. In contrast to the finding in Bönte et al. (2016), Brañas-Garza et al. (2017) do not find a significant relationship between self-reported general risk and right hand 2D:4D (left hand is also insignificant). They do find a significant negative relationship between both left and right hand 2D:4D and risk taking measured by the lottery choice task.

[5] The authors also look at recreational, social, ethical and health risk. They find significance only between recreational and social risk, and 2D:4D, for the Caucasian male subsample only.

[6] In addition, they explore non-linear effects using quartiles of the 2D:4D ratio.

[7] They also look at career risk and find no significant relationship with right hand 2D:4D.

Study	Samp.	Men						Women						Both sexes					
		L	L L^2	R	R R^2	M	M M^2	L	L L^2	R	R R^2	M	M M^2	L	L L^2	R	R R^2	M	M M^2
Dreber and Hoffman (2007), study 1	()													S- (120)		NS (116)			
Dreber and Hoffman (2007), study 2	()													NS (116)		NS (115)			
Apicella et al. (2008)	()	NS (85)		NS (88)															
Sapienza et al. (2009)	()	NS (116)		NS (116)		NS (116)		NS (65)		NS (65)		NS (65)		NS (181)		NS (181)		NS (181)	
Coates and Page (2009)	47			S-															
Brañas-Garza and Rustichini (2011), risk attitude	()			NS (72)						S+ (116)						NS (188)			
Brañas-Garza and Rustichini (2011), combined risk aversion	()			S- (72)						NS (116)						NS (188)			
Garbarino et al. (2011)	151																	S-	
Stenstrom et al. (2011), financial or overall risk^	()			S- (219)						NS (194)									
Sytsma (2014), gain domain*	()	NS (105)		NS (98)		NS (92)		S- (29)		S- (24)		NS (23)		S- (134)		NS (122)		NS (115)	
Sytsma (2014), loss domain*	()	S- (105)		NS (98)		NS (92)		NS (29)		NS (24)		NS (23)		NS (134)		NS (122)		NS (115)	
Sytsma (2014), average*	()	S- (105)		NS (98)		NS (92)		NS (29)		NS (24)		NS (23)		S- (134)		NS (122)		NS (115)	
Aycinena et al. (2014)	()	NS (106)	NS, NS (106)	NS (106)	NS, NS (106)			NS (78)	NS, NS, (78)	NS (78)	NS, NS, (78)			NS (184)	NS, NS (184)	NS (184)	NS, NS (184)		
Drichoutis and Nayga (2015)	138																NS, NS		
Schipper (2014), gains*	()			NS (103)						NS (71)									
Schipper (2014), losses*	()			NS (111)						NS (80)									
Schipper (2014), gains white*	()			NS (47)						NS (25)									
Schipper (2014), losses white*	()			NS (50)						NS (27)									
Schipper (2014), gains asian*	()			NS (52)						NS (41)									
Schipper (2014), losses asian*	()			NS (56)						NS (48)									
Bönte et al. (2016), investment risk^*	432													NS		NS			
Bönte et al. (2016), general risk^*	432													NS		S-			
Barel (2017), general risk^	211													NS		S-			
Barel (2017), financial risk^	211													NS		NS			
Chicaiza-Becerra and Garcia-Molina (2017), full	123													NS		NS			
Chicaiza-Becerra and Garcia-Molina (2017), midland	115													NS		NS			
Brañas-Garza et al. (2017), risk preference	664													S-		S-			
Brañas-Garza et al. (2017), general risk attitude^	704													NS		NS			
Lima de Miranda et al. (2018)	()													NS (144)	NS, NS (144)	NS (145)	NS, NS (145)		
Alonso et al. (2018)	390													NS		NS			

L: left hand. R: right hand. M: mean of left and right hands. NS: not significant. S+: significant positive relationship. S-: significant negative relationship.
*multiple other controls; ^questionnaire elicitation of risk

Table 2: Risk taking studies

1.3 Competitiveness

Whilst there is evidence for gender differences in self-selection into competition (Niederle and Vesterlund (2007)),[8] there exists substantially less literature looking at the relation between prenatal testosterone exposure and willingness to compete, relative to the other economic preferences discussed. Given the gender differences observed in this scenario, the hypothesis tested in the existing literature is that higher testosterone is associated with higher competitiveness, leading to a negative relationship between 2D:4D and the willingness to compete.

Two papers investigate the relationship between 2D:4D and the willingness to compete. Apicella et al. (2011) use self-selection into a competitive tournament or a piece rate scheme for payment in a maze solving task to measure willingness to compete. In an all male sample (83 left hand, 86 right hand), they find no association between left or right hand 2D:4D and the binary willingness to compete measure. Using a similar task, and with larger sample sizes, Bönte et al. (2017), find no significant results for either left or right hand 2D:4D, in two independent mixed-gender samples. They also use a self-reported competitiveness measure in addition to the experimentally elicited measure, and find a significant negative relationship between right hand 2D:4D and self-reported competitiveness in both samples, but not for the left hand.

2 Method

2.1 Experimental procedures and design

The data on 2D:4D were collected in conjunction with a study on the influence of the oral contraceptive pill (Ranehill et al. (2017)). The pre-analysis plan specifying the analysis prior to completion of data collection for this study was posted on the Open Science Framework website on the 21st of August 2015 (available at `http://osf.io/he8nb/`). However, the 2D:4D measure was not part of the main planned analyses in this double-blind randomised study. The exact analyses for the 2D:4D measure were therefore not specified in the pre-analysis plan. Instead it was stated in the pre-analysis plan that the 2D:4D data would be used to carry out tests of previous 2D:4D results reported as statistically significant in the literature (i.e. the data were collected to be able to replicate previous findings). The previously reported results in the literature are therefore the starting point for our analyses, but ideally our tests should have been exactly specified in the pre-analysis plan.

The participants in the study were 340 healthy women aged 18-35 years recruited following the criteria used in the oral contraceptive study.[9] Participants in this study thus had agreed to participate in a randomized controlled trial on the effects of the contraceptive pill. For details on how participants were recruited, the criteria for inclusion and exclusion, and further sample characteristics see Ranehill et al. (2017). Approximately 60% of participants reported an education level of university studies (ongoing) or a university degree. We have no ethnicity data on our sample of participants. The statistical analysis is based on 330 observations as some participants did not complete the data collection (7 discontinued treatment and did not complete the data collection, and 3 had missing hand measurements).

[8] Although there is some evidence that this effect is context dependent (Gneezy et al., 2009)

[9] Such as having a body mass index between 19–30, willing to start using oral contraceptives, being fluent in the Swedish language, being a non-smoker, not being pregnant and so on.

Study	Sample	Men						Women						Both sexes				
		L	L	R	R	M	M	L	L	R	R	M	M	L	L	R	R	M
			L sqr		R sqr		M sqr		L sqr		R sqr		M sqr		L sqr		R sqr	
Apicella et al. (2011)*	()	NS (83)		NS (86)														
Bönte et al. (2017), study 1 behavioural measure	461													NS		NS		
Bönte et al. (2017), study 1 self-reported measure	461													NS		S-		
Bönte et al. (2017), study 2 behavioural measure	150													NS		NS		
Bönte et al. (2017), study 2 self-reported measure	618													NS		S-		

L: left hand. R: right hand. M: mean of left and right hands.
NS: not significant. S+: significant positive relationship. S-: significant negative relationship.
*include a control for sexual orientation

Table 3: Competitiveness studies

The economic experiments on decision making were also reported and analysed in Ranehill et al. (2017). The tests measured dictator game giving, financial risk taking, and willingness to compete. The dictator game giving measure was elicited in a modified dictator game where the participant was asked to allocate SEK 100[10] between herself and a charitable organization, repeated five times with a different charity organisation in each repetition. The average donation across the five decisions is used as our measure of dictator game giving. We include five dictator game decisions in order to reduce measurement error.

We measure risk taking with repeated lottery choices, involving 18 decisions between a certain payoff, and a 50:50 gamble to win a larger amount of money than the safe option, or SEK 0. The certain payoff amounts varied from SEK 40 to 280, and the gamble amounts were either SEK 200, 300 or 400. The percentage of choices of the gamble (i.e. the number of times the gamble was chosen over the certain payoff) is used as our measure of risk taking.

Measuring willingness to compete consisted of asking participants to solve simple tasks of adding numbers for three minutes, first under a non-competitive piece-rate payment scheme of SEK 5 for each correct answer, and then under a competitive tournament payment scheme of SEK 10 for each correct answer only if more tasks were solved than a random competitor (a participant selected from a previous session), otherwise the pay was zero (with SEK 5 for each person in the case of a tie). Then, in the last part, the participant could select to be paid either under the non-competitive piece rate scheme or the competitive tournament scheme. For our willingness to compete measure, we used the choice of competitive tournament scheme in this part (dummy variable where 1 is choice of competitive tournament scheme).

2D:4D results in the literature are sometimes presented for the left hand, sometimes for the right hand, and sometimes for the average of both hands. Following the existing literature, we therefore present results for all these three 2D:4D measures. In the literature results are sometimes presented for a linear model and sometimes a squared term is added to allow for a non-linear relationship. Following the existing literature, we therefore present results both without (the linear model) and with a quadratic term. In total we therefore estimate 18 regression models; 6 models for each outcome measure. In the models with a squared term we evaluate the significance of 2D:4D as the significance of the regression coefficient for the squared 2D:4D, but we also report the significance of an F-test for the joint significance of 2D:4D and the squared 2D:4D.

2.2 Collection of 2D:4D

Digit measurement expressed in millimetres (mm) was performed for digit two (2D) and digit four (4D), using a Vernier digital calliper 0-150 mm (USA, Cocraft) with a precision of 0.01mm. Digit length was directly measured by two raters from the mid-point of the proximal crease of the proximal phalanx to the distal tip of the distal phalanx for 2D and 4D on both left and right hand. The reliability of direct measurement of digits was tested, demonstrating a high repeatability and differences between subjects greater than measurement errors (Savic et al., 2017). The mean value of two measurements of the 2D and 4D length was calculated and then divided to create the 2D:4D ratio, which was used for further statistical analysis.

[10]SEK 100 corresponds to roughly USD 11.

3 Results

Overall we report results for 18 regression variations, with 6 different specifications for the explanatory variables run separately using OLS for the 3 dependent variables, representing our outcome measures of dictator game giving, risk taking, and the willingness to compete. We note that the correlation between left and right hand 2D:4D in our sample is 0.63. Table 4 shows the means and standard deviations for the 2D:4D measures and the outcome variables.

Table 4: Summary Statistics

	mean	std. dev.
Giving	40.748	30.356
Risk	0.550	0.186
Comp.	0.424	0.495
2D:4D LH	0.967	0.033
2D:4D RH	0.980	0.031
2D:4D Avg	0.973	0.029
2D:4D LH sqr	0.935	0.063
2D:4D RH sqr	0.961	0.062
2D:4D Avg sqr	0.948	0.056
Observations	330	

We report the regression results in the following three tables, grouped by outcome measure. Table 5 shows the results for the dictator game giving measure, whilst Table 6 shows risk taking and Table 7 shows the willingness to compete as the dependent variable.

Table 5: Dictator game giving results

	(1)	(2)	(3)	(4)	(5)	(6)
	Giving	Giving	Giving	Giving	Giving	Giving
2D:4D LH	23.2			-930.5		
	(50.87)			(1655.11)		
2D:4D RH		6.50			3729.8	
		(55.70)			(1993.72)	
2D:4D Avg			18.7			2648.6
			(57.81)			(2537.98)
2D:4D LH sqr				491.8		
				(854.80)		
2D:4D RH sqr					-1892.6	
					(1012.69)	
2D:4D Avg sqr						-1350.6
						(1306.45)
Constant	18.3	34.4	22.5	480.1	-1795.0	-1256.6
	(49.20)	(54.62)	(56.29)	(800.98)	(981.03)	(1232.14)
N	330	330	330	330	330	330
F	0.21	0.014	0.11	0.25	1.75	0.63
p	0.65	0.91	0.75	0.78	0.18	0.53

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

Table 6: Risk taking results

	(1) Risk	(2) Risk	(3) Risk	(4) Risk	(5) Risk	(6) Risk
2D:4D LH	0.42			-6.21		
	(0.31)			(11.15)		
2D:4D RH		0.29			27.3	
		(0.34)			(13.89)	
2D:4D Avg			0.44			16.0
			(0.36)			(16.01)
2D:4D LH sqr				3.42		
				(5.75)		
2D:4D RH sqr					-13.7	
					(7.04)	
2D:4D Avg sqr						-7.99
						(8.22)
Constant	0.14	0.27	0.12	3.35	-13.0	-7.45
	(0.30)	(0.34)	(0.35)	(5.41)	(6.85)	(7.79)
N	330	330	330	330	330	330
F	1.83	0.69	1.48	1.13	2.13	1.23
p	0.18	0.41	0.23	0.32	0.12	0.29

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

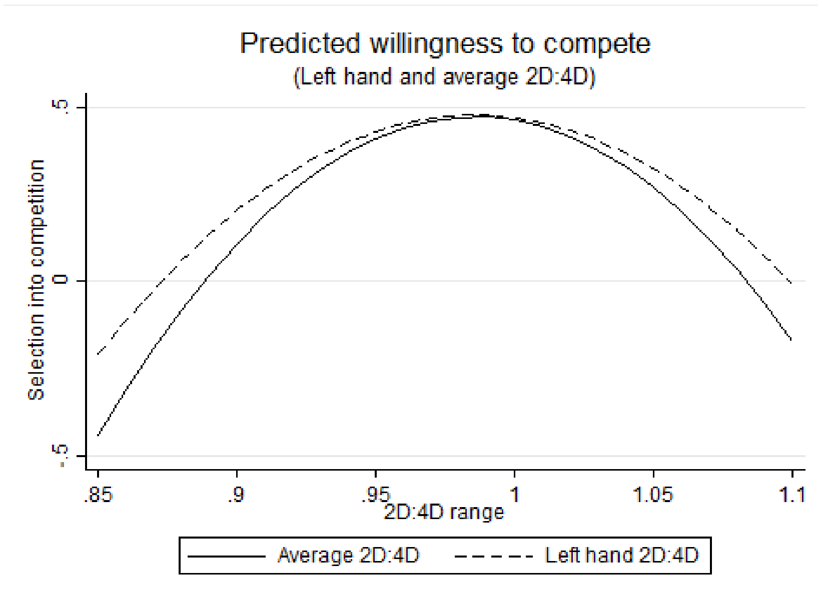
Table 7: Willingness to compete results

	(1) Comp.	(2) Comp.	(3) Comp.	(4) Comp.	(5) Comp.	(6) Comp.
2D:4D LH	1.19			73.4**		
	(0.80)			(22.83)		
2D:4D RH		0.76			43.2	
		(0.86)			(36.33)	
2D:4D Avg			1.21			97.3**
			(0.92)			(33.93)
2D:4D LH sqr				-37.2**		
				(11.84)		
2D:4D RH sqr					-21.6	
					(18.46)	
2D:4D Avg sqr						-49.3**
						(17.44)
Constant	-0.73	-0.32	-0.76	-35.7**	-21.2	-47.5**
	(0.78)	(0.85)	(0.89)	(11.00)	(17.87)	(16.49)
N	330	330	330	330	330	330
F	2.21	0.77	1.74	7.46	1.07	5.22
p	0.14	0.38	0.19	0.00068**	0.34	0.0059*

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.005$

We find no evidence of an effect of 2D:4D on dictator game giving or risk taking ($p > 0.05$). For competitiveness we find no evidence in the linear models either ($p > 0.05$). When we add a squared term we find significant evidence ($p < 0.005$ for the squared 2D:4D coefficient) in both the regression for left hand 2D:4D and competitiveness, and the regression for the average 2D:4D of the two hands and competitiveness. However, these regression specifications are not among those that have previously been reported in the literature for the willingness to compete. We plot the predicted relationships from these significant specifications to illustrate the interpretation of the predicted relationships, using the range of 2D:4D that we see in our data.[11]



The willingness to compete outcomes predicted by our regression equations show an inverse U-shaped relationship where, across a range of 2D:4D values from 0.85 to 1.1, low 2D:4D (synonymous with high prenatal testosterone exposure) predicts low competitiveness, which does not fit with the pre-existing hypothesis that high testosterone correlates with high competitiveness.[12] The highest willingness to compete is instead associated with mid-range 2D:4D for this predicted relationship. If the hypothesis tested in the existing literature were to hold here, we would see a decreasing relationship. As most 2D:4D measurements are below 1, we see that most of the distribution of observations would lie to the left of the peak, in the region of an increasing relationship, which is the opposite to the hypothesised relationship. The estimated inverse U-shaped relationship is thus unlikely to represent a real effect.

4 Discussion

In this study we find little evidence of 2D:4D correlating with economic preferences in a sample of 330 women. The only two significant regression specifications ($p < 0.005$) are not in the hypothesised direction and are not consistent with any previous findings, and are thus likely to be a false positive. The study by Ranehill et al. (2017) that was run in conjunction, but looking at the effect of the oral

[11] Our range of 2D:4D is within that commonly seen in the literature of around $0.8 - 1.2$.

[12] We note that some of the predicted values for the willingness to compete are negative, which is an implication of using OLS for estimation.

contraceptive pill, also did not find any impact of the pill on economic preferences.

Our null results could be due to several reasons. First, 2D:4D may be a reliable proxy of prenatal testosterone exposure but prenatal testosterone exposure may not correlate with economic preferences and previous results are false positive results. Second, 2D:4D may be a reliable proxy of prenatal testosterone but the relation between prenatal testosterone exposure and economic preferences is so weak that with 330 women we do not have sufficient statistical power to detect true positive results. Third, 2D:4D may be a weak or noisy proxy of prenatal testosterone but the relation between prenatal testosterone exposure and economic preferences is actually strong; but again we could then be underpowered to detect true positive results. Fourth, 2D:4D may be a weak or noisy proxy of prenatal testosterone and there is also a weak relation between prenatal testosterone exposure and economic preferences; again we could then be underpowered to detect true positive results. Fifth, 2D:4D may not correlate with economic preferences among women, thus our study would be set up to not find anything since we have only women in our sample. Given previous literature it is not clear to us why this should make a difference but additional high-powered studies, with pre-analysis plans, on men or mixed gender would be useful. In sum, more work is needed to disentangle these five possible explanations for our null results.

In a related vein, the evidence linking sex hormone administration to economic preferences is also inconclusive with most studies failing to reject the null hypothesis of no effect. The few significant findings (as well as the null results) need, however, to be interpreted with caution because of low statistical power and the many researcher degrees of freedom (see the recent review by Dreber and Johannesson (2018) for more information).

In sum, more work is needed with larger sample sizes and pre-registered hypotheses in order to have enough statistical power to find small effects of 2D:4D on economic preferences. Additionally, studies using improved indicators of prenatal testosterone exposure may be warranted.

References

Abreu, Ana Paula and Ursula B Kaiser, "Pubertal development and regulation," *The Lancet Diabetes & Endocrinology*, 2016, *4* (3), 254–264.

Alonso, Judit, Roberto Di Paolo, Giovanni Ponti, and Marcello Sartarelli, "Facts and Misconceptions about 2D: 4D, Social and Risk Preferences," *Frontiers in Behavioral Neuroscience*, 2018, *12*, 22.

Apicella, Coren L, Anna Dreber, Benjamin Campbell, Peter B Gray, Moshe Hoffman, and Anthony C Little, "Testosterone and financial risk preferences," *Evolution and human behavior*, 2008, *29* (6), 384–390.

_ , _ , Peter B Gray, Moshe Hoffman, Anthony C Little, and Benjamin C Campbell, "Androgens and competitiveness in men.," *Journal of Neuroscience, Psychology, and Economics*, 2011, *4* (1), 54.

_ , Victoria A Tobolsky, Frank W Marlowe, and Kathleen W Miller, "Hadza hunter-gatherer men do not have more masculine digit ratios (2D: 4D)," *American journal of physical anthropology*, 2016, *159* (2), 223–232.

Arnold, Arthur P, "The organizational–activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues," *Hormones and behavior*, 2009, *55* (5), 570–578.

Aycinena, Diego, Rimvydas Baltaduonis, and Lucas Rentschler, "Risk preferences and prenatal exposure to sex hormones for ladinos," *PloS one*, 2014, *9* (8), e103332.

Barel, Efrat, "2D: 4D, Optimism, and Risk Taking," *Current Psychology*, 2017, pp. 1–9.

Benjamin, Daniel J, James O Berger, Magnus Johannesson, Brian A Nosek, E-J Wagenmakers, Richard Berk, Kenneth A Bollen, Björn Brembs, Lawrence Brown, Colin Camerer et al., "Redefine statistical significance," *Nature Human Behaviour*, 2018, *2* (1), 6.

Bönte, Werner, Vivien D Procher, and Diemo Urbig, "Biology and selection into entrepreneurship—The relevance of prenatal testosterone exposure," *Entrepreneurship Theory and Practice*, 2016, *40* (5), 1121–1148.

_ , _ , _ , and Martin Voracek, "Digit Ratio (2D: 4D) Predicts Self-Reported Measures of General Competitiveness, but Not Behavior in Economic Experiments," *Frontiers in behavioral neuroscience*, 2017, *11*, 238.

Brañas-Garza, Pablo and Aldo Rustichini, "Organizing effects of testosterone and economic behavior: Not just risk taking," *PloS one*, 2011, *6* (12), e29842.

_ and Jaromir Kovarik, "Digit Ratios and Social Preferences: A Comment on Buser (2012)," 2013.

_ , Jaromír Kovářík, and Levent Neyse, "Second-to-fourth digit ratio has a non-monotonic impact on altruism," *PloS one*, 2013, *8* (4), e60419.

_ , Matteo Galizzi, and Jeroen Nieboer, "Experimental and self-reported measures of risk taking and digit ratio (2D: 4D): evidence from a large, systematic study," *International Economic Review*, 2017.

Brown, Windy M, Melissa Hines, Briony A Fane, and S Marc Breedlove, "Masculinized finger length patterns in human males and females with congenital adrenal hyperplasia," *Hormones and behavior*, 2002, *42* (4), 380–386.

Buser, Thomas, "Digit ratios, the menstrual cycle and social preferences," *Games and Economic Behavior*, 2012, *76* (2), 457–470.

Charness, Gary and Uri Gneezy, "Strong evidence for gender differences in risk taking," *Journal of Economic Behavior & Organization*, 2012, *83* (1), 50–58.

Chicaiza-Becerra, Liliana Alejandra and Mario Garcia-Molina, "Prenatal testosterone predicts financial risk taking: Evidence from Latin America," *Personality and Individual Differences*, 2017, *116*, 32–37.

Coates, John M and Lionel Page, "A note on trader Sharpe Ratios," *PloS one*, 2009, *4* (11), e8036.

Cohen-Bendahan, Celina, *Biological roots of sex differences: A longitudinal twin study*, C. Cohen-Bendahan, 2005.

Croson, Rachel and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature*, 2009, *47* (2), 448–474.

de Miranda, Katharina Lima, Levent Neyse, and Ulrich Schmidt, "Risk preferences and predictions about others: No association with 2D: 4D ratio," *Frontiers in Behavioral Neuroscience*, 2018, *12*, 9.

Dreber, Anna and Magnus Johannesson, "Sex Hormones and Economic Decision Making in the Lab: A Review of the Causal Evidence," in Oliver Schultheiss and Pranjal Mehta, eds., *Routledge International Handbook of Social Neuroendocrinology*, Routledge International Handbooks, Taylor & Francis Group, 2018.

_ **and Moshe Hoffman**, "Portfolio selection in utero," *Stockholm School of Economics*, 2007.

Drichoutis, Andreas C and Rodolfo M Nayga, "Do risk and time preferences have biological roots?," *Southern Economic Journal*, 2015, *82* (1), 235–256.

Eckel, Catherine C and Philip J Grossman, "Men, women and risk aversion: Experimental evidence," *Handbook of experimental economics results*, 2008, *1*, 1061–1073.

Galis, Frietson, Clara MA Ten Broek, Stefan Van Dongen, and Liliane CD Wijnaendts, "Sexual dimorphism in the prenatal digit ratio (2D: 4D)," *Archives of sexual behavior*, 2010, *39* (1), 57–62.

Galizzi, Matteo M and Jeroen Nieboer, "Digit ratio (2D: 4D) and altruism: evidence from a large, multi-ethnic sample," *Frontiers in behavioral neuroscience*, 2015, *9.*

Garbarino, Ellen, Robert Slonim, and Justin Sydnor, "Digit ratios (2D: 4D) as predictors of risky decision making for both sexes," *Journal of Risk and Uncertainty*, 2011, *42* (1), 1–26.

Gelman, Andrew and Eric Loken, "The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time," 2013.

Gneezy, Uri, Kenneth L Leonard, and John A List, "Gender differences in competition: Evidence from a matrilineal and a patriarchal society," *Econometrica*, 2009, *77* (5), 1637–1664.

Grimbos, Teresa, Khytam Dawood, Robert P Burriss, Kenneth J Zucker, and David A Puts, "Sexual orientation and the second to fourth finger length ratio: a meta-analysis in men and women.," *Behavioral Neuroscience*, 2010, *124* (2), 278.

Hiraishi, Kai, Shoko Sasaki, Chizuru Shikishima, and Juko Ando, "The second to fourth digit ratio (2D: 4D) in a Japanese twin sample: Heritability, prenatal hormone transfer, and association with sexual orientation," *Archives of sexual behavior*, 2012, *41* (3), 711–724.

Hönekopp, Johannes and Mirjam Schuster, "A meta-analysis on 2D: 4D and athletic prowess: Substantial relationships but neither hand out-predicts the other," *Personality and Individual Differences*, 2010, *48* (1), 4–10.

_ **and Steven Watson**, "Meta-analysis of the relationship between digit-ratio 2D: 4D and aggression," *Personality and Individual Differences*, 2011, *51* (4), 381–386.

Lombardo, Michael V, Emma Ashwin, Bonnie Auyeung, Bhismadev Chakrabarti, Kevin Taylor, Gerald Hackett, Edward T Bullmore, and Simon Baron-Cohen, "Fetal testosterone influences sexually dimorphic gray matter in the human brain," *Journal of Neuroscience*, 2012, *32* (2), 674–680.

Lutchmaya, Svetlana, Simon Baron-Cohen, Peter Raggatt, Rebecca Knickmeyer, and John T Manning, "2nd to 4th digit ratios, fetal testosterone and estradiol," *Early human development*, 2004, *77* (1), 23–28.

Manning, John T, Diane Scutt, James Wilson, and D Iwan Lewis-Jones, "The ratio of 2nd to 4th digit length: a predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen.," *Human Reproduction (Oxford, England)*, 1998, *13* (11), 3000–3004.

McIntyre, Matthew H, "The use of digit ratios as markers for perinatal androgen action," *Reproductive biology and endocrinology*, 2006, *4* (1), 10.

Medland, Sarah E, John C Loehlin, and Nicholas G Martin, "No effects of prenatal hormone transfer on digit ratio in a large sample of same-and opposite-sex dizygotic twins," *Personality and Individual Differences*, 2008, *44* (5), 1225–1234.

Miller, Edward M, "Prenatal sex hormone transfer: A reason to study opposite-sex twins," *Personality and Individual Differences*, 1994, *17* (4), 511–529.

Millet, Kobe and Siegfried Dewitte, "Second to fourth digit ratio and cooperative behavior," *Biological psychology*, 2006, *71* (1), 111–115.

_ **and** _ , "The presence of aggression cues inverts the relation between digit ratio (2D: 4D) and prosocial behaviour in a dictator game," *British Journal of Psychology*, 2009, *100* (1), 151–162.

Nelson, Julie A, "Are Women Really More Risk-Averse than Men? A Re-Analysis of the Literature Using Expanded Methods," *Journal of Economic Surveys*, 2015, *29* (3), 566–585.

Niederle, Muriel and Lise Vesterlund, "Do women shy away from competition? Do men compete too much?," *The Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.

Phoenix, Charles H, Robert W Goy, Arnold A Gerall, and William C Young, "Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig," *Endocrinology*, 1959, *65* (3), 369–382.

Puts, David A, Michael A McDaniel, Cynthia L Jordan, and S Marc Breedlove, "Spatial ability and prenatal androgens: meta-analyses of congenital adrenal hyperplasia and digit ratio (2D: 4D) studies," *Archives of sexual behavior*, 2008, *37* (1), 100.

Ranehill, Eva, Niklas Zethraeus, Liselott Blomberg, Bo von Schoultz, Angelica Lindén Hirschberg, Magnus Johannesson, and Anna Dreber, "Hormonal Contraceptives Do Not Impact Economic Preferences: Evidence from a Randomized Trial," *Management Science*, 2017.

Sapienza, Paola, Luigi Zingales, and Dario Maestripieri, "Gender differences in financial risk aversion and career choices are affected by testosterone," *Proceedings of the National Academy of Sciences*, 2009, *106* (36), 15268–15273.

Savic, Ivanka, Louise Frisen, Amirhossein Manzouri, Anna Nordenstrom, and Angelica Lindén Hirschberg, "Role of testosterone and Y chromosome genes for the masculinization of the human brain," *Human brain mapping*, 2017, *38* (4), 1801–1814.

Schipper, Burkhard C, "Sex hormones and choice under risk," 2014.

Simmons, Joseph P, Leif D Nelson, and Uri Simonsohn, "False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant," *Psychological science*, 2011, *22* (11), 1359–1366.

Stenstrom, Eric, Gad Saad, Marcelo V Nepomuceno, and Zack Mendenhall, "Testosterone and domain-specific risk: Digit ratios (2D: 4D and rel2) as predictors of recreational, financial, and social risk-taking behaviors," *Personality and Individual Differences*, 2011, *51* (4), 412–416.

Sytsma, Tobias, "Handling Risk: Testosterone and Risk Preference, Evidence from Dhaka, Bangladesh," 2014.

van Anders, Sari M, Philip A Vernon, and Christopher J Wilbur, "Finger-length ratios show evidence of prenatal hormone-transfer between opposite-sex twins," *Hormones and Behavior*, 2006, *49* (3), 315–319.

van de Beek, Cornelieke, Jos HH Thijssen, Peggy T Cohen-Kettenis, Stephanie HM van Goozen, and Jan K Buitelaar, "Relationships between sex hormones assessed in amniotic fluid, and maternal and umbilical cord serum: what is the best source of information to investigate the effects of fetal hormonal exposure?," *Hormones and Behavior*, 2004, *46* (5), 663–669.

Ventura, T, MC Gomes, Ana Pita, MT Neto, and A Taylor, "Digit ratio (2D: 4D) in newborns: influences of prenatal testosterone and maternal environment," *Early Human Development*, 2013, *89* (2), 107–112.

Voracek, Martin and Lisa Mariella Loibl, "Scientometric analysis and bibliography of digit ratio (2D: 4D) research, 1998–2008," *Psychological reports*, 2009, *104* (3), 922–956.

_ **and Stefan G Dressler**, "Digit ratio (2D: 4D) in twins: heritability estimates and evidence for a masculinized trait expression in women from opposite-sex pairs," *Psychological Reports*, 2007, *100* (1), 115–126.

_ **, Jakob Pietschnig, Ingo W Nader, and Stefan Stieger**, "Digit ratio (2D: 4D) and sex-role orientation: Further evidence and meta-analysis," *Personality and Individual Differences*, 2011, *51* (4), 417–422.

Zheng, Zhengui and Martin J Cohn, "Developmental basis of sexually dimorphic digit ratios," *Proceedings of the National Academy of Sciences*, 2011, p. 201108312.