Experimental and self-reported measures of risk taking and digit ratio (2D:4D): Evidence from a large, systematic study

May 31 2017

Pablo Brañas-Garza[a], Matteo M. Galizzi[b]*, Jeroen Nieboer[c]

International Economic Review, forthcoming

Abstract: Using a large (*n=704*) sample of laboratory subjects, we systematically investigate the links between the digit ratio - a biomarker for pre-natal testosterone exposure - and two measures of individual risk taking: (i) risk preferences over lotteries with real monetary incentives, and (ii) self-reported risk attitude. The digit ratio (also called 2D:4D) is the ratio of the length of the index finger to the length of the ring finger, and we consider both hands' digit ratios. Previous studies have found that the digit ratio correlates with risk taking in some subject samples, but not others. In our sample, we find that both the right-hand and the left-hand digit ratio are significantly associated with risk preferences: subjects with lower digit ratios tend to choose riskier lotteries. Neither digit ratio, however, is associated with self-reported risk attitude.

Keywords: Testosterone; 2D:4D ratio; risk preferences; risk attitudes.

JEL codes: C91, C92, D44, D81, D87.

---

[a] Middlesex University London, Economics Department, Business School, The Burroughs, Hendon, London NW4 4BT, United Kingdom.

[b] London School of Economics and Political Science, LSE Behavioural Science Hub, Houghton Street, London WC2A 2AE, United Kingdom. *Corresponding author. E-mail: m.m.galizzi@lse.ac.uk. Phone: +44 (0)20 7955 5386.

[c] London School of Economics and Political Science, Behavioural Research Lab, Houghton Street, London WC2A 2AE, United Kingdom.

1. Introduction

We report findings from a laboratory experiment conducted with a large sample of subjects, which systematically investigates the links between two different measures of individual risk-taking and the *digit ratio* (also known as 2D:4D), the ratio of the length of the index finger to the length of the ring finger. People's digit ratio has been shown to correlate negatively with their pre-natal exposure to androgens, such as the steroid hormone testosterone (Goy and Ewen, 1980; Lutchmaya et al., 2004; Hönekopp et al., 2007; Hönekopp and Watson, 2010; Zheng and Cohn, 2011). Given the difficulty of measuring pre-natal androgen exposure directly, we opt for the digit ratio as a biomarker to investigate how early-life physiology shapes economic behavior in adult life.

Previous digit ratio studies provide evidence that pre-natal testosterone exposure is associated with several types of important economic and financial behavior. Economic experiments show significant correlations between digit ratio and dictator game giving (Branas-Garza et al., 2013; Galizzi and Nieboer, 2015), cognitive reflection (Bosch-Domenech et al., 2014), contributions to a public good (Cecchi and Duchoslav, 2016), overconfidence bias under incentivized conditions (Dalton and Ghosal, 2014; Neyse et al., 2016), and effort provision (Neyse et al., 2014). In the domain of finance, low digit ratio individuals achieve higher trading profits (Coates and Herbert, 2008, Coates et al., 2009), are more likely to self-select into the financial services profession (Sapienza et al., 2009), and are more active and risk-taking traders (Cronqvist et al., 2016). These findings suggest that the preferences underlying these choices - such as people's appetite for competition and risk - are partly determined before birth.

Findings to date strongly suggest a biological basis for economic behavior, complementary to recent research on genetic inheritance of economic behaviors (Rangel et al. 2008; Cesarini et al., 2009; Dreber et al., 2009; Kuhnen and Chiao, 2009; Zhong et al., 2009). Similar to genetic factors, pre-natal hormone exposure can thus shape one's physiology in ways that affect a variety of social and economic outcomes over the life time. The evidence base for the relationship between pre-natal hormones and adult behavior is broad – in both non-human mammals and humans, measures of pre-natal hormones have been shown to correlate with post-natal behavior (Hines, 2006; Hines et al., 2015). Most evidence points to the period from 8 to 24 weeks of fetal gestation as a key stage, during which a marked difference in androgen levels is observed between male and female fetuses (Rodeck et al., 1985; Finegan et al., 1989), leading to different degrees of 'masculinization' of the brain (Manning, 2002).[4] The digit ratio correlates with these androgen levels and is similarly dimorphic – men have lower digit ratios than women. Consequently, much of the literature on pre-natal androgen exposure and digit ratio has focused on correlations with sexually dimorphic behavior, such as athletic achievement (Tester and Campbell, 2007), desire for dominance (Neave et al., 2003), traffic offenses (Schwerdtfeger et al. 2010) and stereotypical childhood play behaviors (Hines, 2006).

Our study focuses on an economic behavior that is often said to be sexually dimorphic: risk taking. Although we study the distribution of risk taking within sexes, we suggest that pre-natal testosterone exposure may contribute to the behavioral observation that women tend to be, on average, more risk averse than men. The latter finding, with important implications in a range of economic situations, has been documented in both experimental and observational economic studies (Byrnes et al., 1999; Croson and Gneezy, 2009). Of course, a multitude of (biological and social) factors may lead to a differentiation between the sexes on risk aversion**, a**nd the observed risk taking tendencies of men and women will overlap to a large extent. However, at least part of the observed difference may have its origins in pre-natal androgen exposure.

We investigate the hypothesis that differences in pre-natal testosterone exposure give rise to different levels of risk aversion, with lower digit ratios being associated with more risk taking. Several prior studies of financial risk taking provide evidence of such a relationship within samples of both sexes (Dreber and Hoffman, 2007; Garbarino et al., 2011) or at least within male sub-samples (Brañas-Garza and Rustichini, 2011; Ronay and von Hippel, 2010; Strenstrom et al., 2011). Our study contributes to the literature on digit ratio and risk taking with a systematic investigation of the relationships between the digit ratios of a large subject sample (*n=704*) and two distinct economic measures of risk taking: (i) revealed risk preferences (RP) over monetary incentives, as measured by the elicitation task developed by Binswanger (1980; 1981; see also Eckel and Grossman, 2002; 2008),

---

[4] This hypothesis is supported by a large body of animal research showing that exposure of the brain's androgen receptors at this stage influences various aspects of brain development (Manning, 2002; Hines, 2006).

and (ii) un-incentivized self-reported risk attitudes (RA), as measured by the Dohmen et al. (2011) scale.

To the best of our knowledge, ours is the first study to date to systematically report, for a large sample of subjects, the associations between both the right-hand digit ratio (R2D:4D) and the left-hand digit ratio (L2D:4D), and two different experimental measures of risk taking, one incentivized and one hypothetical. As explained more in detail in Section 2, our approach to measurement, sample size, and econometric controls for ethnicity, are specifically designed to mitigate some of the issues that may have driven mixed results in the literature to date.

Our main findings are as follows. First, R2D:4D and L2D:4D are significantly negatively correlated with RP: subjects with lower R2D:4D and L2D:4D tend to make riskier choices in the experimental lottery test with real monetary payments. It is worth noting that the negative correlation of the L2D:4D with an experimental measure of risk preferences has not been previously reported by the literature. Second, and in contrast to RP, the R2D:4D and L2D:4D are not significantly associated with RA. In sum, incentivized experimental measures of risk taking correlate with both hands' digit ratios, but hypothetical measures do not.[5]

The rest of the article is structured as follows. Section 2 contains a detailed discussion of the background literature on digit ratio and on its relationship with risk taking. Section 3 describes the methods, while Section 4 presents the results. Section 5 discusses the main findings and concludes.

2. Background.

Digit ratio and pre-natal testosterone exposure

Before we discuss the literature on risk taking, it is worth examining the evidence for the digit ratio as a biomarker for pre-natal testosterone exposure.[6] The 'exposure' is that of the brain's androgen receptors to testosterone - an exposure that is typically much higher for male than female fetuses, since the male fetus produces testosterone in larger amounts (in the Leydig cells of the testes, while females produce it in the

---

[5] This result is in line with Neyse et al. (2014) that show that low digit ratio subjects respond to incentives under real monetary payoffs, but not under hypothetical payoffs.

[6] For an excellent extended summary, see Apicella et al. (2015) or Hines et al. (2015).

adrenal glands near the kidney). Effective exposure may also vary with the hormone levels of the mother (Hines, 2006; Talarovičová et al., 2009). There are four strands of empirical evidence that support the existence of a significant, negative relationship: people with lower digit ratios were exposed to higher levels of pre-natal testosterone.

First, there is *direct evidence from the amniotic fluid*: using a small mixed-sex sample of two-year olds (n=29), Lutchmaya et al. (2004) found that digit ratio is related to testosterone and the testosterone-to-estradiol ratio in utero. Using a larger sample of newborns (n=102), Ventura et al. (2013) found a similar relationship between digit ratio and testosterone in plasma (p=0.04).[7] However, decomposing these results by sex shows significant effects for girls (p=0.03 and p=0.09 for right and left hand digit ratios) but not for boys (both p>0.1). Follow-up research with larger samples seems desirable, as well as studies to fill the evidence gap on the relationship between pre-natal testosterone exposure and digit ratios in adolescent or adult subject samples. There is evidence, however, that digit ratios are stable 3 months after fetal gestation (Galis et al., 2010; Malas et al., 2006) and longitudinally stable in samples of children and adolescents (McIntyre et al., 2005; Trivers et al., 2006).

Second, there is evidence from *androgen spillovers in zygotic twins*: females with a male twin have lower digit ratios than females with a female twin (Van Anders et al., 2006). The channel of influence is a hypothesized ‘hormone-transfer’ between the twins in utero (Miller, 1994), although the support for this theory is somewhat limited.

Third, there is evidence from *individuals with sex hormone-related syndromes*: conditions that limit the production of, or the brain’s sensitivity to, androgens. Subjects with Congenital Adrenal Hyperplasia (CAH) – characterized by increased androgen production - have lower digit ratios than control subjects (Brown et al., 2002). Males with Complete Androgen Insensitivity Syndrome (CAIS) have higher digit ratios than controls (Berenbaum et al., 2009). Similarly, males with Klinefelter’s syndrome – associated with low fetal androgen levels – have higher digit ratios than controls (Manning et al., 2013).

A fourth source of evidence is the *laboratory study of non-human mammals*. Since experimentation with pre-natal testosterone administration on human fetuses is ethically unacceptable, testosterone administration in laboratory animals may be the

---

[7] Note also that the amniotic fluid contains higher levels of testosterone for male fetuses than for female fetuses (Rodeck et al., 1985; Finegan et al., 1989).

closest substitute. Increasing parental testosterone levels in pregnant rats has been found to lead to lower digit ratios of both male and female fetuses (Talarovičová et al., 2009). Similarly, Auger et al. (2013) exposed male rat fetuses to estrogenic and anti-androgenic disruptors and found that this led to higher digit ratios. In mice, testosterone administration in utero leads to lower digit ratios, whereas estrogen administration leads to higher digit ratios (Zheng and Cohn, 2011). Although replications with other species of mammals seem desirable to strengthen the evidence base, we do note that these findings fit into a broad experimental literature that documents the effects of pre-natal testosterone administrations on mammalian brain development (Arnold, 2009; Hines et al., 2015).

Unlike levels of circulating hormones, which may change as a response to an individual's context and actions (see Archer, 2006), the stability of the digit ratio implies it cannot be shaped by the individual's previous behavior. With the issue of two-way causation out of the way, the question remains whether there is any relationship between the digit ratio and circulating testosterone levels. The jury is still out: although Hönekopp et al.'s (2007) meta-analysis on a sizeable body of research did not find any relationship between the digit ratio and circulating sex hormone levels in adults, more recent research suggests that the digit ratio is associated with circulating sex hormones under challenging situations, like fighting or competition (Coates et al., 2010; Crewther et al., 2015).

The previous paragraph hints at a more general question: is the effect of pre-natal hormones on the developing brain the only relevant influence that the digit ratio proxies for? This is currently unclear. Most of the research on digit ratio seems to make a tacit assumption that selection into different levels of testosterone exposure in utero is independent of other indirect influences on behavior. We note that this assumption is untested and may not hold. It is, for example, possible that physiological characteristics of the mother affect both the effectively level of testosterone exposure in utero and aspects of the child's upbringing. Whether this is merely a theoretical possibility or a factor of significance is a topic worthy of further research.[8] We now turn our focus to the relationship between digit ratio and risk taking.

Digit ratio and risk taking

A number of studies – summarized in Table 1 - have explored the relationship between digit ratio and experimental measures for risk taking, yielding mixed evidence to date. In particular:

- Five studies find a negative, significant relationship between digit ratio and risk taking: people with a lower digit ratio take more risk. Dreber and Hoffman (2007) and Garbarino et al. (2011) find this relationship for both males and females, while Ronay and von Hippel (2010), Brañas-Garza and Rustichini (2011) and Strenstrom et al. (2011) find a statistically significant relationship for males only.
- Five studies find a statistically not significant association between digit ratio and risk taking (Apicella et al., 2008; Sapienza et al., 2009; Aycinena et al., 2014; Drichoutis and Nayga, 2015; Schipper, 2015b).

As Table 1 shows, methods differ greatly between studies, both in terms of subject pool and of the measurement of key variables. First, significant relationships appear either in Caucasian samples or male-only samples: not a single significant result is found for females only. This asymmetric effect might be related to the fact that males are exposed, on average, to higher amounts of testosterone in utero.

Second, mixed results in the literature to date may stem from a combination of selective sampling from particular ethnicities and small sample sizes.[9] The studies cited in Table 1 consider either samples of (predominantly) Caucasian subjects (Dreber and Hoffman, 2007; Garbarino et al., 2011; Ronay and von Hippel, 2010; Brañas-Garza and Rustichini, 2011) or relatively small samples of ethnically diverse subjects (Apicella et al., 2008; Sapienza et al., 2009; Drichoutis and Nayga, 2015;

---

[8] Note, for example, the evidence that fetal and maternal testosterone levels are positively correlated (Gitau et al., 2005; but see also Rodeck et al., 1985) and the evidence of a negative correlation between a mother's digit ratio and the likelihood of having a son (Kim et al., 2015).

[9] Ethnicity has been cited as an important source of variation in digit ratio: Manning (2002) and Manning et al. (2014) report that the variation of digit ratio between ethnic groups, and even between Caucasians of different European origin, is larger than the variation between sexes within an ethnic group. Such large variation makes it harder to detect a relationship between digit ratio and risk taking in small samples.

Schipper, 2015b). Weaker relationships between the digit ratio and risk taking in studies with mixed-ethnicity samples might therefore be due to a relationship between digit ratio and risk taking that is mediated by ethnicity. In fact, all the studies reporting significant relationships are conducted with Caucasians.[10] To address any concerns about sample size, we recruit a large sample of subjects (*n=704*) consisting of students of different ethnicities.

Table 1: Summary of the existing studies on 2D:4D ratio and experimental measures for risk taking

	Year	**Exp.**	**Money**	**Measure**	**Hands**	**Ethnicities**	**n_M, n_F**	**Result**
Dreber & Hoffman	2007	GP	Yes	Scanner	Both	Caucasian	87, 65	(-) all
Apicella et al.	2008	GP	Yes	Scanner	Both	Mixed	89, 0	No
Sapienza et al.	2009	HL	Yes	Calliper	Mean	Mixed	117, 66	No
Ronay & Hippel	2010	BART	Yes	Scanner	Mean	Caucasian	52, 0	(-) males
Brañas & Rustichini	2011	HL	Not	Photocopy	Right	Caucasian	72, 116	(-) males
Garbarino et al.	2011	MPL	Yes	Scanner	Mean	Caucasian	87, 65	(-) all
Stenstrom et al.	2011	LTI	Not	Calliper	Mean	Caucasian	130, 109	(-) males
Aycinena et al.	2014	HL	Yes	Scanner	Both	Ladino	125, 94	No
Drichoutis & Nayga	2015	HL	Yes	Ruler	Right	Mixed	46, 92	No
Schipper	2015	HL	Yes	Scanner	Right	Mixed	93, 115	No

Note: *Exp* defines the type of experimental measure to elicit risk-taking: *HL* refers to the Holt-Laury test; *GP* refers to the Gneezy-Potters test; *MPL* to multiple price list tests; *LTI* refers to Likert type items; *BART* to the Balloon Analogue Risk Task. *Mean* refers to the mean of left and right 2D:4D. *Hands* refer to the measure reported in the study. N_M and N_F refer to the number of male and female subjects, respectively. (-) means a statistically significant negative association between 2D:4D and risk taking.

---

[10] Apicella et al. (2008) suggest that the null results found in ethnically diverse samples could be due to small sample sizes: "*If the effect is small, it may not have been detected due to the small sample and possible measurement error associated with calculating 2D:4D.*" (p.388). More generally, the meta-analysis by Hoffman et al. (2013) concludes *"that there is a true relationship between 2D:4D and risk preferences, but because 2D:4D is a noisy measure, we should expect many individual studies to yield null results or even insignificant results in the opposite direction*." (p.13).

Finally, previous studies differ greatly in terms of how the digit ratio measure is taken and subsequently computed. Researchers use various tools (e.g. photocopies, scanners) and then use either the digit ratio of both hands, or the mean digit ratio of the two hands, or the digit ratio of the right hand only. Regarding the R2D:4D, there is some biological evidence to indicate that R2D:4D is more reflective of pre-natal hormone exposure than left-hand digit ratio.[11] The two digit ratio measurements, however, are typically strongly correlated, which may mean that the L2D:4D is simply a noisier measure.

In our study, we follow a standardized procedure to obtain high-quality digit ratio measures from hand scans (Neyse and Brañas-Garza, 2014) and report data on *both* the R2D:4D and the L2D:4D. Note that the actual digit ratio is defined on bone length, something we do not directly observe. Any method that do not use radiographs, therefore, introduces noise into the measurement. The fact that it is only possible to obtain a noisy measure of the digit ratio - which itself is a proxy for pre-natal testosterone exposure – may partly explain the mixture of significant and null results reported thus far. The literature to date may also have been affected by a reporting bias with regards to which gender and which hand is tested for a correlation with risk taking: Apicella et al. (2015), for example, point out that studies that report fewer measures of the digit ratio have a greater proportion of significant results. As with any empirical literature, one cannot rule out the possibility of a 'file drawer' problem (Rosenthal, 1979; Ioannidis, 2005; Simonsohn et al., 2014).

As mentioned, our main contribution concerns the systematic investigation of the relationships between the digit ratios of a large sample and two different economic measures of risk taking. The studies listed in Table 1 use different experimental measures for risk taking, some incentivized with monetary outcomes and some not incentivized. Other studies use self-reported indicators. We collect both incentivized and not incentivized measures of risk taking, and test both for an association with both hands' digit ratios. More in detail:

---

[11] Lutchmaya et al. (2004) find that pre-natal hormone levels are correlated with R2D:4D but not L2D:4D for a sample of 2-year olds; Hönekopp and Watson (2010) find that R2D:4D displays greater variance than L2D:4D between sexes, as well as between healthy people and those affected by Congenital Adrenal Hyperplasia.

- Our first measure is an experimental elicitation task for risk preferences (RP) over real monetary payments developed by Binswanger (1980; 1981) and then applied by Eckel and Grossman (2002; 2008). The RP task involves a choice between six lotteries with different levels of risk. We select this task because its links with the digit ratio have never been previously investigated (Table 1), and because it has the advantage of being simple to understand and intuitive, thus yielding clean and consistent choices (Charness et al., 2013). The RP task, in fact, has already been used to measure risk preferences of large heterogeneous samples of the population (Dave et al., 2010; Galizzi et al., 2016a). The RP task also has drawbacks. For example, compared to the Holt and Laury (2002; 2005) test,[12] the RP task does not allow to discriminate between different degrees of risk seeking, and maps into a rather limited range of constant relative risk aversion (CRRA) parameters that do not directly overlap with the ranges of risk aversion values implied by the standard versions of the Holt and Laury (2002) test (Loomes and Pogrebna, 2014; Crosetto and Filippin, 2015). Nonetheless, a direct systematic comparison of the RP task with the Holt and Laury (2002; 2005) test within a representative sample of the UK population finds a positive and statistically significant correlation between the two measures of risk aversion (Galizzi et al., 2016a).
- Our second measure is a self-reported measure for general risk attitudes (RA) on a 10-point Likert scale developed by Dohmen et al. (2011) which has been introduced in large representative surveys (Josef et al., 2016; Galizzi et al., 2016a) and which has been extensively used in other studies with neurobiological measures (Cesarini et al., 2009; Zethraeus et al., 2009). This procedure also has drawbacks. For example, the procedure does not allow to associate the different individual choices with specific ranges of risk aversion parameters under a CRRA theoretical framework.

Looking at different measures of risk taking is important because risk taking is likely to be a multi-faceted and largely context-specific construct (Jackson et al. 1972; Hershey and Shoemaker, 1980; MacCrimmon and Wehrung, 1990; Viscusi and Evans, 1990; Zeckhauser and Viscusi, 1990; Bleichrodt et al., 1997; Finucane et al.

---

[12] Together with the Charness-Gneezy-Potters method (Gneezy and Potters, 1997; Charness and Gneezy, 2010), the Holt and Laury (2002, 2005) test is one of the other most common procedures to measure risk preferences in experimental economics.

2000; Loewenstein et al. 2001; Weber et al. 2002; Blais and Weber, 2006; Prosser and Wittenberg, 2007; Galizzi et al. 2016b) and because the evidence is mixed on the extent to which different measures correlate and map into each other (see Galizzi et al. 2016a for a summary of the evidence to date on the cross-validity of risk preferences measures). It is thus plausible that incentive-compatible, hypothetical and/or self-reported measures capture different aspects of individual risk-taking (Battalio et al., 1990; Holt and Laury, 2002; 2005; Harrison, 2006). Most of the studies on the links between digit ratio and risk-taking, however, have exclusively looked at Multiple Price List measures such as the already mentioned Holt and Laury (2002) task. Exceptions are the studies by Dreber and Hoffman (2007) and Apicella et al. (2008), who consider the investment task by Gneezy and Potters (2002); Ronay and von Hippel (2010), who use the Balloon Analogue Risk Task (BART) procedure; Brañas-Garza and Rustichini (2011), who use a series of non-incentivized binary lottery choices (including the Holt and Laury, 2002 task); and Stenstrom et al. (2011), who use a questionnaire. As mentioned above, no study to date has ever looked at the links between the digit ratios and risk preferences as measured by the Binswanger (1980; 1981) and Eckel and Grossman (2002; 2008) and the Dohmen et al. (2011) procedures.

3. Methods

All experimental sessions were run at the Behavioural Research Lab (BRL) at the London School of Economics and Political Science (LSE), London. The first round of data collection took place in February and March 2014 (yielding 543 observations); a supplementary round of data collection took place in April 2015 (yielding a further 161 observations). The procedures followed in both rounds were identical. The experimental protocol was approved by the LSE Research Ethics Committee. Subjects were recruited from the BRL mailing list of volunteers (about 5,000 subjects, mostly current and former students of the LSE). There was no other eligibility or exclusion criterion to select subjects. In the email invitation, subjects were not informed about the exact nature of the experiment that would be conducted, and were only told that the experiment would last about an hour, that they would receive £10 as a show-up fee, and that they would have the chance to get an extra payment related to some of the tasks. Subjects could sign up to any of five one-hour sessions starting every hour between 10 am and 5 pm at every working day in the week.

A total of 921 subjects participated in our experimental sessions. Upon arrival, subjects were identified anonymously using an ID code assigned by the online recruitment system (SONA), asked to read an informed consent form, and to sign the latter if they agreed to carry on with the experiment. After the experiment, subjects were led to a separate room where they were presented with a second consent form, which asked for consent to have both of their hands scanned by a high-resolution scanner. Subjects were clearly briefed that participation in this stage was entirely voluntary. A total of 704 subjects gave consent for their hands to be scanned and yielded resulting scans of sufficiently high quality. We thus focus our analysis on these 704 subjects (76.43% of the original sample). Note that this is an underestimation of the actual consent rate, as we lost a number of observations due to a technical issue with the scanner.[13]

We distinguish between risk preferences (RP), subjects' observed choice between monetary *lotteries* which are played out and paid for real at the end of experiment; and risk attitudes (RA), a self-reported measure of risk taking. Both measures were obtained in a computerized questionnaire administered at the start of the experimental session. The questionnaire also contained other items, such as questions about personality and demographic data. The computerized questionnaire was programmed and implemented using Z-Tree (Fischbacher, 2007).

The RP elicitation task we used was the lottery choice originally proposed by Binswanger (1980; 1981) and further applied by Eckel and Grossman (2002; 2008). The task required the subjects to choose between six lotteries with an equal chance of receiving a *low* or *high* cash payment:

- A: low = £28, high = £28;
- B: low = £24, high = £36;
- C: low = £20, high = £44;
- D: low = £16, high = £52;

---

[13] To check for any selection bias of subjects with different characteristics into having their hands scanned, we compared the risk preferences and risk attitudes of subjects who did or did not have their hands scanned. For both the RA measure and RP measure we cannot reject the null hypothesis that the mean of the two samples come from the same distribution (two-tailed Mann-Whitney U test, z=-0.984, p=0.325, for RA, and z=0.757, p=0.449 for RP).

- E: low = £12, high = £60;
- F: low = £2, high = £70.

These choices were thus increasing in the variance of the outcomes and in the risk they represented, with A being the safe bet (a variance of zero) and F being the highest-risk choice (a variance of $\sigma_F^2 = 1156$). To make a choice, subjects clicked one of six radio buttons on their screen, which were labeled with the lottery probabilities and outcomes. Our RP measure thus increases with an individual's appetite for risk. As mentioned, the Binswanger (1980; 1981) and Eckel and Grossman (2002; 2008) procedure has drawbacks: in particular, it does not allow to discriminate among different degrees of risk seeking. Under the assumptions of Expected Utility Theory and that the respondents have a Constant Relative Risk Aversion utility function, in fact, the ranges of CRRA values implied by each lottery choice are [3.46; ∞] for lottery A, [1.16; 3.46] for lottery B, [0.71; 1.16] for lottery C, [0.499; 0.71] for lottery D, [0; 0.499] for lottery E, and [-∞; 0] for lottery F.

The RA measure was the self-reported measure from Dohmen et al. (2011). Each subject was asked the following: "*Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?*". To select an answer between 0 and 10, subjects clicked a radio button on their screen, on which the value 0 was labeled as "*Unwilling to take risks*" and the value 10 was labeled as "*Fully prepared to take risks*". In the on-screen instructions it was made clear to subjects that the question was about their own assessment of their general attitude towards risk. Our RA measure thus increases with individual self-reported risk taking, with values between 0 and 10.

The RA question was asked first, followed by the RP task a few screens later, with the two questions being separated by other questionnaire items unrelated to risk. This separation was designed to avoid subjects, consciously or unconsciously, adjusting their answer to the RP item to match their answer to the RA item. Furthermore, the RP question was preceded by an on-screen announcement that the upcoming choices would affect subjects' earnings. Note that the RP item was followed by several other incentivized decisions - subjects were informed that each of these decisions would have an equal probability of being randomly selected to be played and paid out for real at the end of the experiment. Average earnings per subject for the entire experiment, composed of the £10 show-up fee and potential extra earnings from the incentivized choices, were £19.48. Subjects were paid their earnings in cash at the end of the session.

After the questionnaire and a completely unrelated task, subjects were led into a separate room where the experimenters had set up a computer with a high-resolution scanner (300 DPI on a Canon LIDE 110). Subjects were told: "*Before you leave the laboratory today, we would like to ask you to participate in an optional task. Please can you read the following consent form to see what it involves?*"[14] Subjects were then given time to read an informed consent form, which explained that they would be asked to place both of their hands on a scanner to obtain the digit ratio, which *"…has been shown in various scientific studies to correlate with people's behaviour in the laboratory."* They were reminded that placing their hands on the scanner was completely voluntary and that the data would remain strictly anonymous and confidential (*"…we will not be able to share your digit ratio with anyone, including you"*). Finally, they were told that they could ask as many questions as they wanted.[15]

After the experimental sessions were completed, we recruited two research assistants to provide us with independent measures of the length of the second and fourth finger of each hand.[16] We calculated the digit ratios from the finger length measures and checked the correlation between the digit ratios implied by the measurements from the two research assistants. These correlations (0.895 for left hand, 0.867 for right hand) suggest that measurement was highly accurate. To obtain a single measure of the digit ratio of each hand for our analysis, we computed the average of the two research assistants' ratios (Neyse and Brañas-Garza, 2014).

---

[14] See Appendix B for these instructions.

[15] When subjects asked what kind of behavior the digit ratio predicted, or what the purpose of our study was, the experimenters replied that we were looking for correlations with their answers to the questionnaire that was administered earlier.

[16] The research assistants were told to take as much time as they needed to provide us with reliable measures. Both research assistants used Adobe Photoshop to measure the length of the fingers on the scans. They were instructed by the same experimenter to follow the procedures described in Neyse and Brañas-Garza (2014). The assistants were also given a copy of this procedure, for reference. The two research assistants did not know or meet each other and worked independently at different times. Research assistants had no access to the details of the subjects' whose fingers they were measuring.

4. Results

Summary statistics

Our sample consists of 704 student subjects. The sample consists predominantly of female students (478, 67.89% of the sample). The sample, moreover, is highly ethnically diverse: 244 subjects described themselves as Chinese (34.65% of the sample), 241 as White (34.23%), 96 as South Asian (13.63%), 30 as Black (4.26%) and 93 as 'Other'. Note the relatively small number of Black subjects. Given the composite nature of the 'Other' ethnicities in our sample, we do not report summary statistics for this group of subjects. The subjects are predominantly females in all the ethnic groups: 65.56% of White, 72.13% of Chinese, 75% of South Asian, and 50% of Black subjects are females.

Digit ratios

Table 2 summarizes our left-hand (L2D:4D) and right-hand digit ratio (R2D:4D) measures, in aggregate, and by sex and ethnicity-specific subsamples. Figure 1A shows the sample distribution of the L2D:4D for male and female subjects separately; Figure 1B shows the same for R2D:4D.

Overall, both the L2D:4D and the R2D:4D of the male subjects are lower than those of female subjects. The average R2D:4D is 0.9584 (SD=0.0305) for male subjects and 0.9770 (SD=0.0325) for female subjects; the average L2D:4D is 0.9599 (SD=0.0353) for male subjects and 0.9733 (SD=0.0321) for female subjects. Both differences are strongly statistically significant (p=0.0000).

The significant differences of digit ratios across sexes also hold when the analysis is replicated at ethnicity level, with the only exception of Black subjects. Chinese males have significantly lower L2D:4D and R2D:4D than Chinese females (p=0.0086 and p=0.0002, respectively), and the same holds for White subjects (p=0.0283 and p=0.0007) and for South Asian subjects (p=0.0468 and p=0.0055).

Whilst the difference in digit ratio between sexes is significant, differences between ethnicities are not clear-cut in our sample. In general, the L2D:4D is 0.9661 (SD=0.0296) for Chinese subjects, 0.9720 (SD=0.0329) for White subjects, 0.9733 (SD=0.0370) for South Asians and 0.9650 (SD=0.0476) for Black subjects: the L2D:4D for Chinese subjects are statistically different from the L2D:4D of White subjects (p=0.0156). In general, the R2D:4D is 0.9679 (SD=0.0305) for Chinese subjects, 0.9738 (SD=0.0331) for White subjects, 0.9753 (SD=0.0349) for South Asians, and 0.9595 (SD=0.0352) for Black subjects: the R2D:4D for Chinese and Black subjects are statistically different from the R2D:4D of White subjects (p=0.0184 and p=0.0543, respectively).

Table 2. Summary statistics for Left-Hand and Right-Hand Digit Ratios.

		Left-Hand (L2D:4D)			**Right-Hand (R2D:4D)**		
	Obs.	Mean		St. Dev.	Mean		St. Dev.
All	704	0.9689		0.0337	0.9714		0.0329
Female	478	0.9733	]***	0.0321	0.9770	]***	0.0325
Male	226	0.9599		0.0353	0.9594		0.0305
Chinese	244	0.9661	]**	0.0296	0.9679	]**	0.0305
White	241	0.9720		0.0329	0.9738		0.0331
S-Asian	96	0.9733		0.0370	0.9753		0.0349
Black	30	0.9650		0.0476	0.9595		0.0352

Note: Significant differences between sub-samples (two-tailed Wilcoxon rank-sum test) are shown as brackets in the last column: * = p < 0.1, ** = p < 0.05, *** = p < 0.01.

For males, we found no statistically significant differences in L2D:4D or R2D:4D between ethnic groups. Within the female sub-sample, the differences between the L2D:4D and R2D:4D for White females (0.9752 and 0.9790, respectively) and for Chinese females (0.9694 and 0.9724) are both statistically significant (p=0.0252 and p=0.0230, respectively). Similarly, the differences between the L2D:4D and R2D:4D for White females (0.9752 and 0.9790, respectively) and for Black females (0.9605 and 0.9607) are both statistically significant (p=0.0658 and p=0.0245, respectively).

Risk taking

Only 35 subjects in our sample chose lottery F in the RP task (18 male subjects, and 17 female subjects). Moreover, the ordered probit models (discussed in detail in the next regression analysis subsection) suggested that the estimated threshold parameters for the cut-off points corresponding to the lottery choices E and F were not statistically significantly different from each other, suggesting that the two categories should better be collapsed into the same category. We have therefore recoded the responses to the RP experimental test into 5 categories, taking value 1 if subjects chose the safe lottery A, value 2 if subjects chose lottery B, and so on increasing in risk-seeking, up to value 5 if the subjects chose either lottery E or F.[17]

Table 3. Summary statistics for Risk Preferences and self-reported Risk Attitudes.

		Risk Preferences (RP)		**Risk Attitudes (RA)**	
	Obs.	Mean	St. Dev.	Mean	St. Dev.
All	704	2.794	1.306	4.697	2.273
Female	478	2.714	1.251	4.541	2.238
		] **		] ***	
Male	226	2.971	1.407	5.026	2.315
Chinese	244	2.608	1.349	4.319	2.257
		] **		] ***	
White	241	2.872	1.349	4.971	2.257
S-Asian	96	2.869	1.215	4.697	2.113
Black	30	3.011	1.385	5.433	2.487

Note: Significant differences between sub-samples (two-tailed Wilcoxon rank-sum test) are shown as brackets in the last column: * = $p < 0.1$, ** = $p < 0.05$, *** = $p < 0.01$.

---

[17] We are grateful to an anonymous reviewer for having suggested this analysis. We have also replicated all the estimations of the ordered probit models with six (instead of five) ordered values for the dependent variable (i.e. with choices of lotteries E and F considered in two distinct categories), or only focusing on choices of lotteries A to E, and in all cases we have obtained substantially identical results concerning the associations (or lack of associations) between the digit ratios and the two measures of risk-taking (all available on request).

The left side of Table 3 summarizes our recoded RP measure. The mean value for RP in our sample is 2.794 (SD=1.306). Male subjects in our sample chose riskier lotteries on average, with a mean choice of 2.971 (SD=1.407) compared to 2.714 (SD=1.251) for female subjects, a difference that is statistically significant (p=0.0282). This result is in line with the commonly reported finding that women are more risk averse than men (Charness and Gneezy, 2012; Croson and Gneezy, 2009; Eckel and Grossman, 2008).[18]

With the exception of the Chinese subjects who are significantly more risk averse than the White subjects (p=0.0156), and of Chinese female subjects who are marginally more risk averse than White female subjects (p=0.0913), we find no significant differences between the RP of different ethnicities, neither for the whole sample nor for sex-specific subsamples. Moreover, when looking at each ethnicity separately, we cannot find any statistically significant differences in the RP between sexes.

The right side of Table 3 summarizes our data for the RA measure. The mean value for RA in our sample is 4.697 (SD=2.273). Also according to this measure, male subjects appears slightly more risk seeking, describing themselves as 5.026 on average (SD=2.315) compared to 4.541 (SD=2.238) among female subjects, a difference which is statistically significant (p=0.0087). Risk attitudes among South Asian (4.697) and Black (5.433) subjects are not statistically significantly different from White subjects (4.971), but Chinese subjects (4.319) report to take significantly less risk than White subjects (p=0.0012). None of the differences in risk attitudes are significant considering the subsample of males only, while Chinese females (4.159) report to take significantly less risk than White females (4.892, p=0.0053). Moreover, when looking at each ethnicity separately, we cannot find any statistically significant differences in the RA between sexes.

Figures 1C and 1D report the sample distributions of the responses of male and female subjects to the RP and RA tasks, respectively.

As it can be seen in Figure 1C, male subjects in our sample tend to take more risks than female subjects in the RP task. The figure visually confirms the above mentioned

---

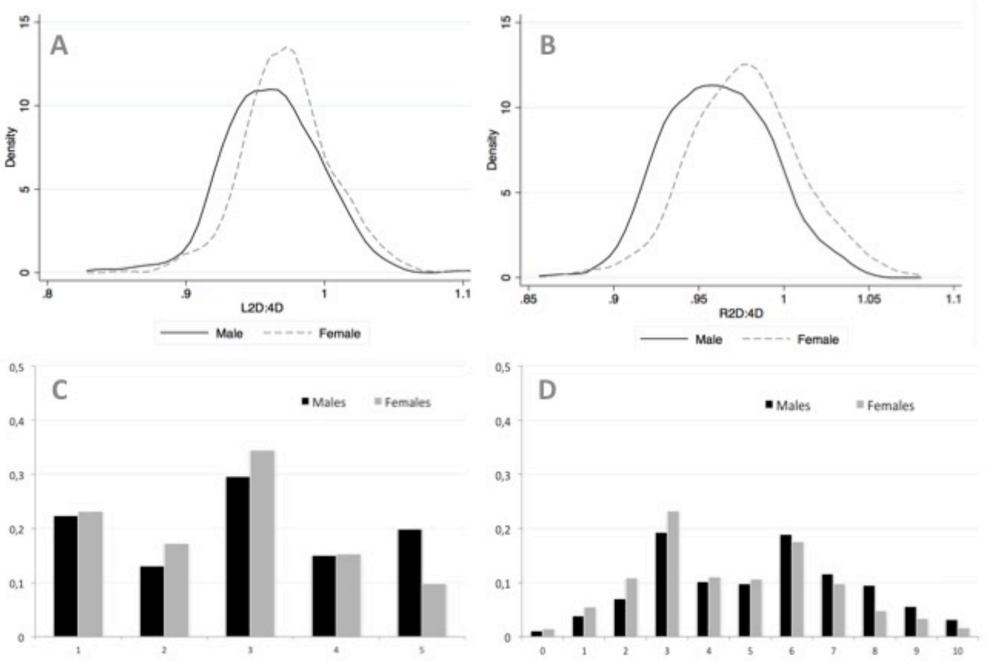
[18] Note that evidence on the difference between male and female risk-taking in the laboratory is currently disputed (see, for instance, Filippin and Crosetto, 2016).

finding that women tend to be more risk averse than men (Charness and Gneezy, 2012; Croson and Gneezy, 2009; Eckel and Grossman, 2008).

Figure 1D shows that, compared to female respondents, male subjects report to be more willing to risk in the RA task (see also Table 3, right side).

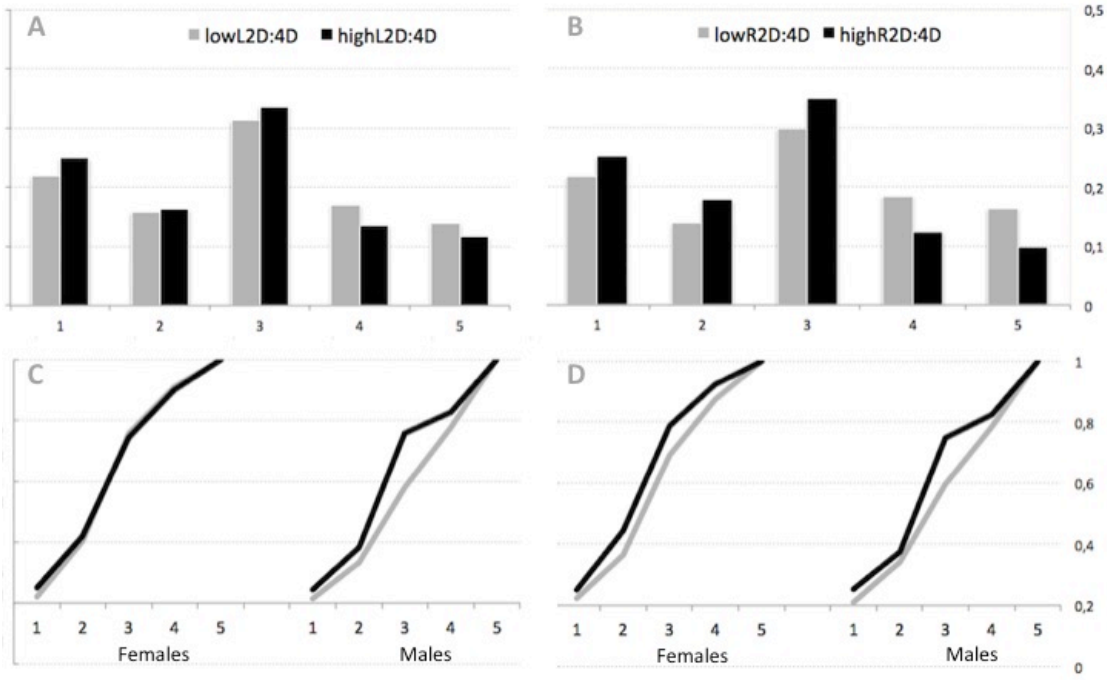
Figure 2A (2B) reports the sample distribution of the responses to the RP task split by low and high L2D:4D (R2D:4D). In particular, the respondents are divided according to whether their L2D:4D (R2D:4D) is below (*Abovemedian=0*) or above (*Abovemedian=1*) the median value of the L2D:4D (R2D:4D) in our sample. Figures 2C and 2D report the corresponding sample distributions of the RP responses split by subject sex.

Figure 1. Panel A: Kernel density for L2D:4D for female (dashed grey line) and male (solid black line) subjects. Panel A: Kernel density for R2D:4D for female (dashed grey line) and male (solid black line) subjects. Panel C: Histogram of Risk Preferences (RP) for female (grey) and males (black) subjects. Panel D: Histogram of Risk Attitudes (RA) for female (grey) and males (black) subjects.



As it can be seen in Figure 2 (right panel), subjects with low R2D:4D (below the median value) tend to take more risks in the RP task than subjects with high R2D:4D (above the median value). The bottom part the Figure 2 focuses on male and female respondents separately. Looking at the cumulative distributions, it can be seen that lottery choices by subjects with digit ratios below the median (both males and females) are first order stochastically dominated by the choices of subjects with digit ratios above the median, which implies that the formers take more risk than the latters.

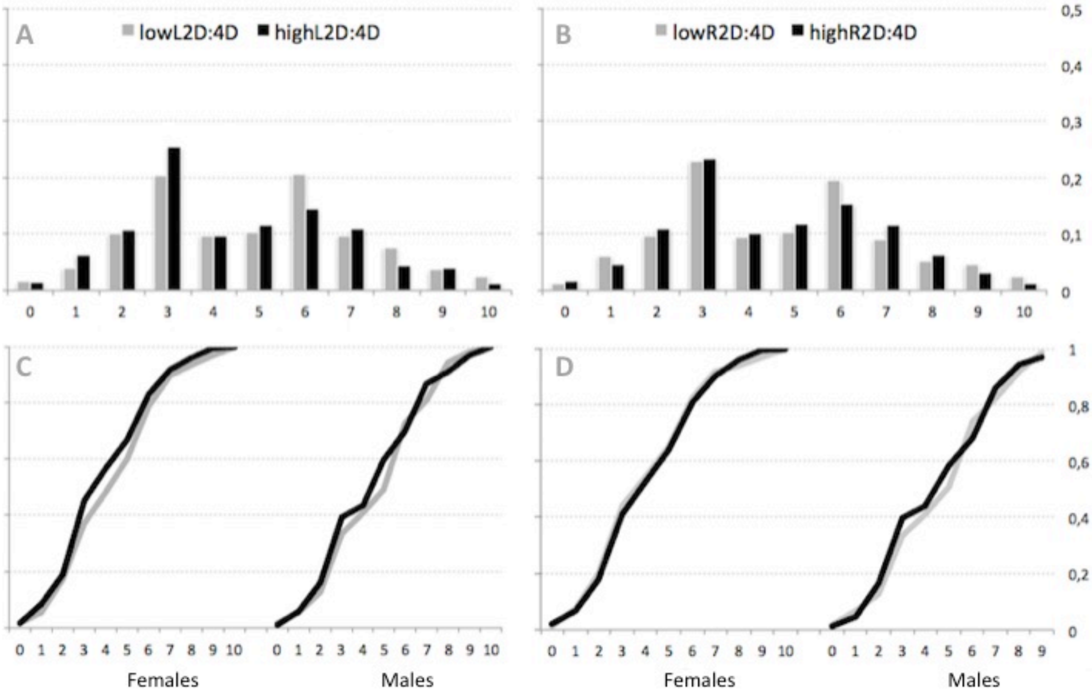
Figure 2. Panel A: Histogram of Risk Preferences (RP) for low (grey) and high (black) LD2:4D subjects, with low (high) LD2:4D referring to values below (above) the median. Panel B: Histogram of Risk Preferences (RP) for low (grey) and high (black) RD2:4D subjects, with low (high) RD2:4D referring to values below (above) the median. Panel C: Cumulative Distribution Functions (CDFs) of Risk Preferences (RP) for low (grey) and high (black) L2D:4D female (left side) and male (right side) subjects, with low (high) LD2:4D referring to values below (above) the median. Panel D: Cumulative Distribution Functions (CDFs) of Risk Preferences (RP) for low (grey) and high (black) R2D:4D female (left side) and male (right side) subjects, with low (high) RD2:4D referring to values below (above) the median.



An analogous pattern emerges when the RP responses are split by below and above the median L2D:4D (Figure 2, left panel), but the difference in the distribution of RP responses is less evident than the analogous difference for the R2D:4D. While we observe strong differences for males, the same pattern is not observed for females (bottom left).

Figure 3A (3B) reports the sample distribution of the responses to the RA task split by low and high L2D:4D (R2D:4D) for both male and female subjects.

Figure 3: Panel A: Histogram of Risk Attitudes (RA) for low (grey) and high (black) LD2:4D subjects, with low (high) LD2:4D referring to values below (above) the median. Panel B: Histogram of Risk Attitudes (RA) for low (grey) and high (black) RD2:4D subjects, with low (high) RD2:4D referring to values below (above) the median. Panel C: Cumulative Distribution Functions (CDFs) of Risk Attitudes (RA) for low (grey) and high (black) L2D:4D female (left side) and male (right side) subjects, with low (high) LD2:4D referring to values below (above) the median. Panel D: Cumulative Distribution Functions (CDFs) of Risk Attitudes (RA) for low (grey) and high (black) R2D:4D female (left side) and male (right side) subjects, with low (high) RD2:4D referring to values below (above) the median.



The respondents are divided again according to whether their high L2D:4D (R2D:4D) is below (*Abovemedian=0*) or above (*Abovemedian=1*) the median value of the L2D:4D (R2D:4D) in our sample. The corresponding sample distributions of the RA responses by low and high L2D:4D and R2D:4D for the male and female subjects are shown below (Figures 3C and 3D, respectively).

As can be seen in Figure 3 (notably panels 3B and 3D), in the RA task there are some differences in the willingness to take risks between the subjects with low R2D:4D (below the median value) and the subjects with high R2D:4D (above the median value): subjects with low R2D:4D (right) seemingly report to be somewhat more willing to take risks. The difference in the distributions of the RA responses, however, is far less evident than the analogous difference in the distributions of the RP responses.

Correlation analysis

Table 4 reports pairwise correlations among the main variables of interest.[19] We first note that, in our sample, L2D:4D and R2D:4D are strongly positively correlated (0.719, p=0.000). Next, looking at the measures of risk taking, we find a significant positive correlation between the incentive-compatible risk preference test and the self-reported risk attitude measure (p=0.000). However, we note that the correlation coefficient is rather low (0.204), in line with other evidence of moderate correlations between the two methods (Crosetto and Filippin, 2015; Galizzi et al., 2016a). This may indicate that self-reported risk attitudes (RA) and risk preferences revealed through experimental tasks with real monetary incentives (RP) capture different aspects of individual risk taking.

Furthermore, the correlation analysis reveals interesting patterns of association between digit ratios and our risk-taking measures. On the one hand, there is a negative and significant correlation between RP and R2D:4D: -0.126 (p=0.001). So, the higher is the R2D:4D - that is, the lower the pre-natal testosterone exposure - the less likely are the subjects to take risk in an incentivized experimental test. The association of RP with L2D:4D is also negative (-0.108) and statistically significant (p=0.005). The

---

[19] For each pairwise correlation, Table 4 also reports the p-value of the t-test of the null hypothesis that the Pearson correlation coefficient between any two given variables is equal to zero.

sign of the association is in line with the existing literature (Dreber et al., 2007; Garbarino et al., 2011; and also Ronay and von Hippel, 2010 and Brañas-Garza and Rustichini, 2011, although for males only).

Table 4: Pairwise correlations between the main variables.

	R2D:4D	L2D:4D	RP	RA
R2D:4D	1			
L2D:4D	0.719*** (0.000)			
RP	-0.126*** (0.001)	-0.108*** (0.005)		
RA	-0.010 (0.792)	-0.021 (0.582)	0.204*** (0.000)	1

Note: Standard errors in parentheses. * = p < 0.1, ** = p < 0.05, *** = p < 0.01.

On the other hand, the self-reported RA measure does not exhibit significant correlations with either digit ratio: although the association is negative with both the L2D:4D (-0.021) and the R2D:4D (-0.010), neither of these are statistically significant (p=0.582 and p=0.792, respectively).

Similar patterns of association hold when only the subsample of male or female subjects is considered (Table A1 and A2 in the Appendix). Notice, however, that while the correlation of RP with RA, of L2D:4D with R2D:4D, and of RP with R2D:4D are all statistically significant for the sex-specific subsamples, the negative association between RP and L2D:4D is not significant in the all-female subsample, and the negative association between RP and R2D:4D is only marginally significant in the all-male subsample.[20]

<u>*Regression analysis*</u>

Digit ratio and Risk Preferences

We also conduct regression analysis to explore the links between digit ratio and risk taking, controlling for sex and ethnicity. We first look at Risk Preferences (RP),

---

[20] The latter result could point to differences between male and female subjects in our sample and/or to differences in the sample size of the two genders subsamples (fewer male subjects).

which we investigate using an Ordered Probit (OP) model. In our OP model, the dependent variable can take five values, from 1 (choosing lottery A) to 5 (choosing either lottery E or F), increasing with individual risk seeking. We first look at sex and ethnicity as explanatory variables and then add digit ratio variables (R2D:4D or L2D:4D) into the OP regressions, retaining controls for sex and ethnicity. Unless stated otherwise, all regression models are conducted pooling all data together and with adjustments to the variance-covariance matrix for possible heteroskedasticity and serial correlation.

Starting with the regressions of RP on individual characteristics, results show (Table A3, Appendix) that female subjects are more risk averse (p=0.024), even when controlling for ethnicity (p=0.032). There is no significant effect for any ethnicity, apart from the Chinese group, with Chinese subjects being significantly more risk averse (p=0.035 and p=0.048 when controlling for sex).

We now turn to the regression models with digit ratio variables, starting with R2D:4D (Table 5) and then replicating with L2D:4D (Table 6). We first look at the R2D:4D as the main explanatory variable for RP, and then add sex, an interaction term between sex and R2D:4D, and ethnicity variables as control variables, while retaining R2D:4D. These regressions are conducted with adjustments to the variance-covariance matrix to account for possible heteroskedasticity and serial correlation.

Table 5 shows that, when included in the regression on its own, the R2D:4D is negatively and strongly significantly associated with RP (p=0.001): subjects with lower R2D:4D tend to be less risk averse, a result which is closely in line with previous studies and with the descriptive and correlation analyses. Importantly, the association of RP with R2D:4D remains statistically significant even when directly controlling for sex (p=0.007), sex and a sex-R2D:4D interaction term (p=0.066), ethnicity (p=0.000), and both sex and ethnicity simultaneously (p=0.002): individuals with lower R2D:4D tend to make less risk-averse choices in the incentive-compatible experimental test. There are no significant sex or sex-R2D:4D interaction effects in the estimations with R2D:4D.[21]

---

[21] The ordered probit estimations in Tables 5 and 6 also show that the threshold parameters for RP appear to be statistically significantly different from each other, suggesting that the five RP categories should not be further collapsed into fewer categories. As already mentioned, we have also replicated all the estimations of the ordered probit models only focusing on choices of lotteries A to E, or with six ordered values for the dependent variable (i.e. with choices of lotteries E and F considered in two

Table 5: RP, R2D:4D and individual characteristics: all subjects (OProbit).

RP	***m1***	***m2***	***m3***	***m4***	***m5***
R2D:4D	-3.939***	-3.420***	-4.658*	-4.457***	-4.003***
	(1.227)	(1.264)	(2.533)	(1.233)	(1.278)
Female		-0.148	-1.820		-0.128
		(0.095)	(2.805)		(0.097)
R2D:4D*Female			1.734		
			(2.907)		
Chinese				-0.248**	-0.237**
				(0.101)	(0.101)
S-asian				-0.001	0.012
				(0.124)	(0.124)
Black				-0.175	-0.189
				(0.217)	(0.220)
Other				0.119	0.116
				(0.140)	(0.140)
Observations	664	664	664	664	664
Pseudo R^2	0.0050	0.0062	0.0064	0.0102	0.0111

Note: Standard errors in parentheses; * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors of the cut-points have been omitted.

Next, we turn to the regression model with L2D:4D. Table 6 shows that, when included in the regression on its own, the L2D:4D is negatively and strongly significantly associated with RP (p=0.006): subjects with lower L2D:4D tend to be less risk averse, a result which is closely in line with the descriptive and correlation analyses and that has never been previously documented in the literature.

The association of RP with the L2D:4D is marginally lower and less statistically significant than with the R2D:4D. Importantly, however, the association of RP with L2D:4D remains statistically significant even when directly controlling for sex (p=0.018), sex and a sex-L2D:4D interaction term (p=0.005), ethnicity (p=0.003), and both sex and ethnicity simultaneously (p=0.009): individuals with lower L2D:4D tend to make less risk-averse choices in the incentive-compatible experimental test. There is a marginally significant sex effect (female subjects tend to make more risk-averse

---

distinct categories), and in all cases we have obtained substantially identical results concerning the associations between the digit ratios and the RP measure. The ordered probit models, however, suggested that the estimated threshold parameters for the cut-off points corresponding to the lottery choices E and F were not statistically significantly different from each other, suggesting that the two categories should better be collapsed into one category.

choices) and a marginally significant sex-L2D:4D interaction effect in the estimations with L2D:4D.

Table 6: RP, L2D:4D and individual characteristics: all subjects (OProbit).

RP	***m1***	***m2***	***m3***	***m4***	***m5***
L2D:4D	-3.354***	-2.923**	-6.186***	-3.645***	-3.249***
	(1.215)	(1.233)	(2.200)	(1.220)	(1.242)
Female		-0.171*	-4.961*		-0.157*
		(0.094)	(2.555)		(0.095)
L2D:4D*Female			4.968*		
			(2.648)		
Chinese				-0.238**	-0.226**
				(0.101)	(0.101)
S-asian				0.001	0.016
				(0.123)	(0.124)
Black				-0.145	-0.166
				(0.212)	(0.217)
Other				0.096	0.095
				(0.141)	(0.141)
Observations	664	664	664	664	664
Pseudo R^2	0.0037	0.0055	0.0072	0.0083	0.0098

Note: Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors of the cut-points have been omitted.

Digit ratio and self-reported Risk Attitudes

We next consider the relationship between Digit Ratio and Risk Attitudes (RA), modelled using an Ordered Probit (OP) model. In our OP model, the dependent variable can take eleven values, associated with the eleven degrees of risk taking that the subjects could self-report. Again, we first conduct a set of regressions with sex and ethnicity as explanatory variables, while the second set of regression models adds the digit ratios (L2D:4D and R2D:4D) retaining controls for sex and ethnicity. Also these regression models are conducted with adjustments to the variance-covariance matrix for possible heteroskedasticity and serial correlation.

Table A4 (Appendix) reports the findings from the OP regression models of RA without digit ratio variables. Female subjects in our sample report significantly lower willingness to take risks. This is in line with what found by Josef et al. (2016) and Galizzi et al. (2016) in representative samples in Germany and the UK, respectively. Furthermore, among the various ethnic groups, only the Chinese subjects report significantly more risk-averse attitudes when directly asked how risk-seeking they are. Both effects are robust to controlling for both sex and ethnicity together.

We now turn to the regression models with digit ratio variables, starting with R2D:4D (Table 7) and then replicating with L2D:4D (Table A5). We first look at the R2D:4D (or L2D:4D) as the main explanatory variable for RA, and then add sex, an interaction term between sex and R2D:4D, and ethnicity dummies as control variables, while retaining R2D:4D (or L2D:4D).

Table 7: RA, R2D:4D and individual characteristics: all subjects (OProbit).

RA	***m1***	***m2***	***m3***	***m4***	***m5***
R2D:4D	-0.370	0.413	-0.819	-0.523	0.168
	(1.127)	(1.162)	(2.036)	(1.129)	(1.165)
Female		-0.220***	-1.904		-0.194**
		(0.085)	(2.391)		(0.086)
R2D:4D*Female			1.746		
			(2.477)		
Chinese				-0.291***	-0.275***
				(0.091)	(0.091)
S-asian				-0.106	-0.089
				(0.115)	(0.115)
Black				0.193	0.174
				(0.211)	(0.207)
Other				-0.101	-0.109
				(0.135)	(0.134)
Observations	704	704	704	704	704
Pseudo R^2	0.0000	0.0023	0.0024	0.0043	0.0060

Note: Standard errors in parentheses; * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors for the cut-points have been omitted.

Next, we turn to the association between RA and R2D:4D, shown in Table 7. In no regression is the R2D:4D significantly associated with self-reported risk attitude, neither on its own, nor when included together with sex and/or ethnicity variables. The only variables significantly associated to RA seem to be again the dummies for female and Chinese subjects, both of whom self-report more risk-averse attitudes.

Table A5 (Appendix) reports the OP models of RA and L2D:4D. As with R2D:4D, there is no significant association between RA and L2D:4D, neither when included in the regressions on its own, nor with sex and/or ethnicity as control variables. Also in the regressions with the L2D:4D, female and Chinese subjects self-report to be more risk averse.[22]

Consistency of results across subsamples

Furthermore, in Tables A6-A9 in the Appendix we also report the results of the estimations obtained in the sub-samples of male and female subjects. For the sake of comparability, for each dependent variable (RP or RA), we report the estimations for the full sample and for the two sex-specific sub-samples in terms of the models where the only explanatory variables are the digit ratios (L2D:4D or R2D:4D) as well as of the models adding the controls for the ethnicity groups. As it can be seen in Tables A6-A9, and in line with the previously reported analysis, in the full sample both the L2D:4D and the R2D:4D are negatively and significantly associated with RP, while none of them is associated with RA.

All the associations are robust to the inclusion of the ethnicity controls, with Chinese being the only ethnic group significantly (negatively) associated with RP. In the male sub-sample, both the L2D:4D and the R2D:4D are negatively and significantly associated with RP, but the association with the R2D:4D is only marginally significant ($p=0.073$). In the female sub-sample, the R2D:4D is negatively and significantly associated with RP, but the L2D:4D is not significantly associated with RP ($p=0.407$). In both the male and the female sub-samples, there is no association between the digit ratios and RA.

---

[22] The ordered probit estimations in Tables 7 and A5 show that the threshold parameters for RA appear sometimes not to be statistically significantly different from each other, suggesting that the eleven RA categories could be collapsed into fewer categories. We have therefore replicated the analysis using ordered probit models with less granular categories and found substantially identical results concerning the lack of significant associations between the digit ratios and the RA measures (all available on request).

Finally, note that all our results are qualitatively identical when the regressions are conducted excluding the respondents in the Black or Other ethnic groups; using OLS models or ordered logit hierarchical regressions; using interval regression models for the ranges of the coefficient of relative risk aversion implied by the different choices in the RP; using standardized z-values for the digit ratios (as done in Garbarino et al., 2011); using the average digit ratio of the two hands instead of the R2D:4D and L2D:4D separately (results not reported but all available on request).

As a further robustness check, note that our results are not affected by corrections for multiple testing. For example, if we adjust the p-values of our pairwise correlation coefficients for R2D:4D and L2D:4D using a conservative correction - such as the Bonferroni (1935) correction - that assumes no correlation between outcome variables, our findings remain substantially unchanged (Table A10): the digit ratios are significantly negatively associated with risk taking in the experimental task, and not significantly associated with self-reported risk attitudes. Less conservative adjustments which allow for correlations among the variables - such as the corrections proposed by Holm (1979), Hochberg (1988), Hommel (1988), Benjamini and Hochberg (1995), or Benjamini and Yekutieli (2001), for example - would *a fortiori* yield the same substantial findings.

5. Discussion and conclusions

To the best of our knowledge, ours is the first study to date to systematically report, for a large ethnically diverse pool of subjects, the associations between one incentivized and one hypothetical measure of risk taking, and both the Right-Hand Digit Ratio (R2D:4D) and the Left-Hand Digit Ratio (L2D:4D).

We report two main findings. First, both the R2D:4D and the L2D:4D are significantly associated with risk preferences measured by an incentive-compatible experimental task: subjects with lower R2D:4D and L2D:4D tend to make significantly riskier choices in the experimental lottery test with real monetary payments. This finding is robust across a wide range of alternative specifications, which vary the estimation strategies, and include sex and ethnicity dummies as well as other controls. We thus contribute to the existing literature (Dreber et al., 2007; Garbarino et al., 2011; Ronay and von Hippel, 2010; Brañas-Garza and Rustichini,

2011) by showing that the association between R2D:4D and financial risk-taking that these studies report for relatively small samples of Caucasian subjects also holds within large samples of ethnically diverse subjects.

Although marginally weaker than the association with the R2D:4D, the association of the L2D:4D with an experimental measure of risk preferences has not been previously reported by the literature. This confirms the importance of separately considering both hands' measures when looking at the links between digit ratios and behavioral attitudes (Dreber and Hoffman, 2007; Apicella et al., 2008).

Second, in contrast to our findings on revealed risk preferences, neither the R2D:4D nor the L2D:4D is significantly associated with risk attitudes measured by a hypothetical question. That is, while incentivized experimental measures of risk taking are related to both hands' digit ratios, hypothetical measures are not. Although our study is the first to test such a relationship for digit ratio and risk taking, this result is in line with the abundant experimental literature showing that self-reported and incentive-compatible measures for economic preferences correlate only imperfectly (Battalio et al., 1990; Blackburn et al., 1994; Cummings et al., 1995, 1997; Rutstrom, 1998; Camerer and Hogarth, 1999; List, 2001; Holt and Laury, 2002, 2005; Harrison, 2006; Lusk and Shogren, 2007).[23]

It is worth reflecting on how our negative findings on hypothetical risk-taking decisions fit into the broader literature. Our finding is in line with the idea that risk-taking is a complex, multi-dimensional aspect of individual behavior, and that different measures could well capture different nuances and angles of risk taking (Jackson et al., 1972; Hershey and Shoemaker, 1980; MacCrimmon and Wehrung, 1990; Viscusi and Evans, 1990; Zeckhauser and Viscusi, 1990; Bleichrodt et al., 1997; Finucane et al. 2000; Weber et al. 2002; Blais and Weber, 2006; Prosser and Wittenberg, 2007; Galizzi et al., 2016a,b). It seems also plausible that one's general tendency to take risk is much more influenced by one's social environment, socio-economic situation, knowledge, and other factors, instead of traits associated with pre-natal hormone exposure. A monetary gamble in a laboratory experiment, a much more instantaneous decision that comes with its own context, does show a correlation with

---

[23] Note also the recent study by Neyse et al. (2016) that shows that low digit ratio subjects were more likely to exhibit overconfidence, but this only held for hypothetical rewards: once real payoffs were used, the evidence of overconfidence of low digit ratio subjects was not statistically significant.

pre-natal hormones. Self-reporting bias aside, the differences between the two risk-taking measures are myriad and it may well be that they rely on different cognitive and physiological processes. A more constructive take is that it may be possible to classify certain decisions under risk as more ‘visceral’ or ‘hormonal’ than others, perhaps shedding some light on the emotional determinants of risk-taking (Loewenstein, 1996; Damasio et al. 2006; Le Doux, 1998; Loewenstein et al. 2001). An alternative interpretation of our results is that there may be a correlation between people’s general tendency to take risk and pre-natal hormones in the population, but that idiosyncrasies of our sample of university students do not allow us to detect this relationship.

Of course, our findings also have limitations. Although both the R2D:4D and the L2D:4D are significantly associated with experimental measures of risk preferences, the digit ratios explain only a very small part of the variance in individual risk taking. This finding is consistent with the remarks of Apicella et al. (2008) on the small size of the digit ratio ‘effect’. We note also the potential for measurement to introduce further noise into the equation.

Another limitation of our study is that it looks at the links between risk taking and digit ratios among subjects in an ethnically diverse, but socially homogeneous, large pool of subjects. It is widely known that university students may be a peculiar and unrepresentative sub-sample of the population (Enis et al., 1972; Gachter et al., 2004; Exadaktylos et al., 2013). Further research is needed to systematically explore the association of digit ratios and risk taking in more socially and culturally diverse groups and in representative samples of the population.

One line of inquiry that deserves further attention is the role of mediating factors. For instance, risk taking has previously been shown to correlate positively with cognitive ability (Frederick, 2005; Benjamin et al., 2013; Dohmen et al., 2010), and so has digit ratio. Bosch-Domenech et al. (2014) find that low 2D:4D males and females score higher in the cognitive reflection test; Branas-Garza and Rustichini (2011) find the same relationship for males’ performance in Raven matrices. Other examples of possible mediating factors are preferences for competition, sensation seeking, optimism, and overconfidence. Disentangling the relationship between pre-natal hormones and various aspects of ability and preference is likely to be a complex task, but one that could greatly enhance our understanding of how personalities are shaped in utero. It may, for example, shed more light on why individuals with low 2D:4D are more likely to self-select into the financial services profession (Sapienza et al., 2009) and are more successful in highly competitive professions like financial trading (Coates and Herbert, 2008).

In closing, it is worth reiterating that, although the digit ratio is relatively easy to measure, and it cannot be altered or manipulated, it is only a proxy. Although the evidence on its association with economic behavior, notably risk taking, is rapidly accumulating, it still leaves us several steps removed from actually measuring the effect of pre-natal hormone exposure. As we discussed earlier, further research is needed on whether pre-natal factors affecting the digit ratio are linked to third factors that may shape one's behavior in later life. An example, which may be of sufficient interest to researchers in its own right, is the relationship between pre-natal hormone exposure, parental hormone levels, and the infant's upbringing. More generally, longitudinal research that links directly observed pre-natal hormone levels to behavior in later life, beyond infancy, would do much to enrich the interpretation of digit ratio studies. With the right kind of tools and sufficiently large subject samples, there is much promise in linking biological and behavioral economic measures.

References

Apicella, C.L., Dreber, A., Campbell, B., Gray, P.B., Hoffman, M., Little, A.C. (2008). Testosterone and financial risk preferences. *Evolution and Human Behavior*, 29, 384-390.

Apicella, C.L., Carré, J.M., Dreber, A. (2015). Testosterone and economic risk taking: A Review. *Adaptive Human Behavior and Physiology*, 1, 358-385.

Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews*, 30(3), 319-345.

Arnold A.P., Breedlove S.M. (1985). Organizational and activational effects of sex steroids on brain and behavior: a reanalysis. *Hormones and Behavior*, 19(4), 469-498.

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

Auger, J., Le Denmat, D., Berges, R., Doridot, L., Salmon, B., Canivenc-Lavier, M. C., Eustache, F. (2013). Environmental levels of oestrogenic and antiandrogenic compounds feminize digit ratios in male rats and their unexposed male progeny. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1768).

Aycinena, D., Baltaduonis, R., Rentschler, L. (2014). Risk preferences and prenatal exposure to sex hormones for Ladinos. *PLoS ONE*, 9(8), e103332.

Battalio, R.C., Kagel, J.C., Jiranyakul, K. (1990). Testing between alternative models of choice under uncertainty: some initial results. *Journal of Risk and Uncertainty*, 3, 25-50.

Benjamin, D.J., Brown, S. A., Shapiro, J.M. (2013). Who is ‘behavioral’? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*, 11(6), 1231-1255.

Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, 57(1), 289-300.

Benjamini, Y., Yekutieli, D. (2001). The control of the false discovery rate in multiple test under dependency. *Annals of Statistics*, 29(4), 1165-1188.

Berenbaum, S. A., Bryk, K. K., Nowak, N., Quigley, C. A., Moffat, S. (2009). Fingers as a marker of prenatal androgen exposure. *Endocrinology*, 150(11), 5119-5124.

Binswanger, H.P. (1980). Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics*, 62, 395-407.

Binswanger, H.P. (1981). Attitudes toward risk: Theoretical implications of an experiment in rural India. *The Economic Journal*, 91 (364), 867-890.

Blackburn, M., Harrison, G.W., Rutstrom, E. (1994). Statistical bias functions and informative hypothetical surveys. *American Journal of Agricultural Economics*, 76(5), 1084-1088.

Blais, A.R., Weber, E.U. (2006). A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgment and Decision Making*, 1(1), 33-47.

Bleichrodt, H., Wakker, P.P., Johannesson, M. (1997). Characterizing QALYs by risk neutrality. *Journal of Risk and Uncertainty*, 15, 107-114.

Bonferroni, C.E. (1935). *Il calcolo delle assicurazioni su gruppi di teste*. Tipografia del Senato.

Bosch-Domènech, A., Brañas-Garza, P., Espín, A. (2014). Can exposure to prenatal sex hormones (2D:4D) predict cognitive reflection? *Psychoneuroendocrinology* 43, 1-10.

Brañas-Garza, P., Rustichini, A. (2011). Organizing effects of testosterone on economic behavior: Not just risk taking. *PLoS ONE*, 6(12), e0029842.

Brañas-Garza, P., Kovarik, J., Neyse, L. (2013). Second-to-fourth digit ratio has a non-monotonic impact on altruism. *PLoS ONE* 8(4): e60419.

Brown, W.M., Hines, M., Fane, B.A., Breedlove, S.M. (2002). Masculinized finger length patterns in human males and females with congenital adrenal hyperplasia. *Hormones and Behavior*, 42(4), 380–386.

Byrnes, J.P., Miller, D.C., Schafer W.D. (1999). Gender differences in risk-taking: a meta-analysis. *Psychological Bulletin*, 125, 367–383.

Cecchi, F., Duchoslav, I. (2016). Prenatal Trauma and Cooperation. Evidence from a Public Goods Game in Post-Conflict Uganda. *Mimeo*.

Cesarini, D., Dawes, C.T., Johannesson, M., Lichtenstein, P., Wallace, B. (2009). Genetic Variation in Preferences for Giving and Risk-Taking. *The Quarterly Journal of Economics*, 124, 809–842.

Charness, G., Gneezy, U. (2010). Portfolio choices and risk attitudes: an experiment. *Economic Inquiry*, 48(1), 133-146.

Charness, G., Gneezy, U., Imas, A. (2013). Experimental Methods: Eliciting Risk Preferences. *Journal of Economic Behavior & Organization*, 87, 43-51.

Coates J.M., Herbert, J. (2008) Endogenous steroids and financial risk taking on a London trading floor. *Proceedings of the National Academy of Sciences of USA*, 105, 6167–6172.

Coates, J.M., Gurnell, M., Rustichini, A. (2009). Second-to-fourth digit ratio predicts success among high-frequency financial traders. *Proceedings of the National Academy of Sciences*, 106(2), 623-628.

Coates, J.M., Gurnell, M., Sarnyai, Z. (2010). From molecule to market: steroid hormones and financial risk-taking. *Philosophical Transactions of the Royal Society B*, 365: 331-343.

Crewther, B., Cook, C., Kilduff, L., Manning, J. (2015). Digit ratio (2D:4D) and salivary testosterone, oestradiol and cortisol levels under challenge. *Early Human Development*, 91, 451-456.

Cronqvist, H., Previtero, A., Siegel, S., White, R.T. (2016). The fetal origins hypothesis in finance: prenatal environment and financial risk taking. *Review of Financial Studies*, 29, 739-786.

Crosetto, P., Filippin, A. (2015). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, 1-29.

Croson, R., Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, 47(2), 448-74.

Cummings R.G., Harrison, G.W., Rutstrom, E.E. (1995). Homegrown values and hypothetical surveys: is the dichotomous choice approach incentive-compatible? *The American Economic Review*, 85, 260-266.

Cummings, R.G., Elliott, S., Harrison, G.W., Murphy, J. (1997). Are hypothetical referenda incentive compatible? *Journal of Political Economy*, 105(3), 609-621.

Dalton, P.S., Ghosal, S. (2014). Self-confidence, overconfidence and prenatal testosterone exposure: evidence from the lab. *Mimeo*. Tilburg University.

Damasio, A.R. (2006). *Descartes' error*. London: Random House.

Dave, C., Eckel, C., Johnson, C., Rojas, C. 2010. Eliciting Risk Preferences: When is Simple Better? *Journal of Risk and Uncertainty*, 41, 219-43.

Dohmen, T., Falk, A., Huffman, D., Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *The American Economic Review*, 100(3), 1238-1260.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G. (2011). Individual risk attitudes: measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522-50.

Dreber, A., Hoffman M (2007). Portfolio selection in utero. *Mimeo*. Stockholm School of Economics.

Dreber, A., Apicella, C.L., Eisenberg, D.T.A., Garcia, J.R., Zamoree, R.S., Lume, J.K., Campbell, B. (2009). The 7R polymorphism in the dopamine receptor D4 gene (DRD4) is associated with financial risk taking in men. *Evolution and Human Behavior*, 30, 85-92.

Drichoutis, A.C., Nayga, R.M. Jr. (2015). Do risk and time preferences have biological roots? *Southern Economic Journal* 82(1), 235-256.

Eckel, C.C., Grossman, P.J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior*, 23(4), 281-95.

Eckel, C.C., Grossman, P.J. (2008). Men, Women and Risk Aversion: Experimental Evidence. Handbook of Experimental Economics Results vol 1, 1061-1073.

Enis, B.E., Cox, K., Stafford, J. (1972). Students as subjects in consumer behavior experiments. *Journal of Marketing Research*, 6, 72-74.

Exadaktylos, F., Espín, A.M., Brañas-Garza, P. (2013). Experimental subjects are not different. *Scientific Reports* 3, 1213.

Filippin, A., Crosetto, P. (2016). A reconsideration of gender differences in risk attitudes. *Management Science,* forthcoming.

Finegan, J.K., Bartleman, B., Wong, P.Y. (1989). A window for the study of prenatal sex hormone influences on postnatal development. *The Journal of Genetic Psychology*, 150, 101–112.

Finucane, M.L., Alhakami, A., Slovic, P., Johnson, S.M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, 13(1), 1-17.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19(4), 25-42.

Gachter, S., Herrmann, B., Thoni, C. (2004). Trust, voluntary cooperation and socio-economic background: survey and experimental evidence. *Journal of Economic Behavior & Organization*, 55, 505-531.

Galis, F., Ten Broek, C.M.A., Van Dongen, S., Wijnaendts, L.C.D. (2010). Sexual Dimorphism in the Prenatal Digit Ratio (2D:4D). *Archives of Sexual Behavior* 39(1): 57-62.

Galizzi, M.M., Machado, S., Miniaci, R. (2016a). Temporal stability, cross-validity, and external validity of risk preferences measures: experimental evidence from a UK representative sample. *Mimeo*, London School of Economics: http://eprints.lse.ac.uk/67554/.

Galizzi, M.M., Miraldo, M., Stavropoulou, C. (2016b). In sickness but not in wealth: Field evidence on patients' risk preferences in the financial and health domain. *Medical Decision Making*, 36(4), 503-517.

Galizzi, M.M., Nieboer, J. (2015). Digit ratio and altruism: evidence from a large, multi-ethnic sample. *Frontiers in Behavioral Neuroscience*, 9:41, doi: 10.3389/fnbeh.2015.00041.

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), 1-26.

Gitau, R., Adams, D., Fisk, N.M., Glover, V. (2005). Fetal plasma testosterone correlates positively with cortisol. *Archives of Disease in Children. Fetal and Neonatal Edition,* 90, F166–F169.

Gneezy, U., Potters, J. (1997). An Experiment on Risk Taking and Evaluation Periods. *The Quarterly Journal of Economics*, 112(2), 631-45.

Goy, R.W., McEwen, B.S. (1980). *Sexual differentiation of the brain*. Cambridge, MA: MIT Press.

Harrison, G.W. (2006). Hypothetical bias over uncertain outcomes. In J. List (Ed.) *Using Experimental Methods in Environmental and Resource Economics*. Northampton, MA: Elgar.

Hershey, J.C., Schoemaker, P.J.H. (1980). Risk Taking and Problem Context in the Domain of Losses: An Expected Utility Analysis. *Journal of Risk and Insurance*, 47(1), 111-32.

Hines, M. (2006) Prenatal testosterone and gender-related behaviour. *European Journal of Endocrinology* 155, 115–121.

Hines, M., Constantinescu, M. and D. Spencer (2015). Early androgen exposure and human gender development. *Biology of Sex Differences* 6(3), 23-44.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4), 800-802.

Hoffman, M., Jordan, J., Yoeli, E. (2013). The evolutionary basis of sex differences in Risk Preferences. *Mimeo*. University of California.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65-70. Holt, C.A., Laury, S.K. (2002). Risk aversion and incentive effects. *The American Economic Review*, 92(5), 1644-55.

Holt, C.A., Laury, S.K. (2005). Risk aversion and incentive effects: New data without order effects. *The American Economic Review*, 95(3), 902-904.

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, 75(2), 383-386.

Hönekopp, J., Bartholdt, L., Beier, L., Liebert, A. (2007). Second to fourth digit length ratio (2D:4D) and adult sex hormone levels: new data and a meta-analytical review. *Psychoneuroendocrinology*, 32, 313-321.

Hönekopp, J., Watson, S. (2010). Meta-analysis of digit ratio 2D:4D shows greater sex difference in the right hand. *American Journal of Human Biology*, 22(5), 619-30.

Hönekopp, J. (2011). Relationships between digit ratio 2D:4D and self-reported aggression and risk taking in an online study. *Personality and Individual Differences*, 51(1), 77–80.

Ioannidis, J.P.A. (2005). Why most published research findings are false. *PLOS Medicine*, 2, 696-701.

Jackson, D.N., Hourany, L., Vidmar, N.J. (1972). A four-dimensional interpretation of risk taking. *Journal of Personality*, 40(3), 483-501.

Josef, A.K., Richter D., Samanez-Larkin, G.R., Wagner, G.G., Hertwig, R., Mata, R. (2016). Stability and change in risk-taking propensity across the adult life span. *Journal of Personality and Social Psychology*, 111(3), 430-450.

Kim, T.B., Oh, J.K., Kim, K.T., Yoon, S.J., Kim, S.W. (2015). Does the Mother or Father Determine the Offspring Sex Ratio? Investigating the Relationship between Maternal Digit Ratio and Offspring Sex Ratio. *PLoS ONE*, 10(11), e0143054.

Kuhnen, C.M., Chiao, J.Y. (2009). Genetic determinants of financial risk taking. *PLoS ONE*, 4(2), e4362. doi:10.1371/journal.pone.0004362

Lea, S.E.G., Webley, P. (2006). Money as a tool, money as a drug: the biological psychology of a strong incentive. *Behavioural and Brain Sciences*, 29(2), 161-209.

LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. New York: Simon and Schuster.

Loewenstein, G.F. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 65, 272-92.

Loewenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127(2), 267-286.

Loomes, G., Pogrebna, G. (2014). Measuring Individual Risk Attitudes when Preferences are Imprecise. *The Economic Journal*, 124, 569–593.

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. doi: 10.1016/j.earlhumdev.2003.12.002

Lusk, J.K., Shogren, J.F. (2007). *Experimental Auctions: Methods and Applications in Economics and Business Research*. Cambridge: Cambridge University Press.

MacCrimmon, K.R., Wehrung, D.A. (1990). Characteristics of risk taking executives. *Management Science*, 36(4), 422-435.

Malas M.A., Dogan S., Evcil E.H., Desdicioglu K. (2006). Fetal development of the hand, digits and digits ratio (2D:4D). *Early Human Development,* 82, 469–475.

Manning, J.T. (2002). *Digit ratio: a pointer to fertility, behavior, and health*. New Brunswick, NJ: Rutgers University Press.

Manning, J.T., Fink, B., Trivers, R. (2014). Digit ratio (2D:4D) and gender inequalities across nations. *Evolutionary Psychology*, 12(4), 757-768.

Manning, J.T., Kilduff, L.P., Trivers, R. (2013). Digit ratio (2D:4D) in Klinefelter's syndrome. *Andrology*, 1(1), 94–99.

Manning, J., Kilduff, L., Cook, C., Crewther, B., & Fink, B. (2015). Digit ratio (2D:4D): A biomarker of prenatal sex steroids and adult sex steroids in challenge situations. *Frontiers in Endocrinology*, 5(9), 3398.

McIntyre, M.H., Ellison, P.T., Lieberman, D.E., Demerath, E., Towne, B. (2005). The development of sex differences in digital formula from infancy in the Fels Longitudinal Study. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1571), 1473-1479.

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

Neave, N., Laing, S., Fink, B., Manning, J.T. (2003). Second to fourth digit ratio, testosterone and perceived male dominance. *Proceedings of the Royal Society B: Biological Sciences*, 270(1529), 2167–2172.

Neyse, L., Brañas-Garza, P. (2014). Digit ratio measurement guide, *MPRA Working Paper* 54134.

Neyse, L., Friedl, A., Schmidt, U. (2014). Payment Schemes, Effort Provision and Prenatal Testosterone Exposure. *Mimeo.*

Neyse, L., Bosworth, S., Ring, P., Schmidt, U. (2016). Overconfidence, Incentives and Digit Ratio. *Scientific Reports* 6, 23294.

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

Prosser, L.A., Wittenberg, E. (2007). Do risk attitudes differ across domains and respondent types? *Medical Decision Making*, 27(3), 281-287.

Rangel. A., Camerer, C., Montague, P.R. (2008). A Framework for studying the neurobiology of value-based decision making. *Nature Reviews: Neuroscience*, 9, 1-13.

Rodeck, C.H., Gill, D., Rosenberg, D.A., Collins, W.P. (1985) Testosterone levels in midtrimester maternal and fetal plasma and amniotic fluid. *Prenatal Diagnosis,* 5, 175–181.

Ronay, R., von Hippel, W. (2010). Power, testosterone, and risk ‐ taking. *Journal of Behavioral Decision Making*, *23*(5), 473-482.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638-641.

Sapienza, P., Zingales, L., Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences*, 106(36), 15268-15273.

Schwerdtfeger, A., Heims, R., Heer, J. (2010). Digit ratio (2D:4D) is associated with traffic violations for male frequent car drivers. *Accident Analysis & Prevention*, 42, 269-274.

Schipper, B.C. (2015a). Sex hormones and competitive bidding. *Management Science*, 61, 249-266.

Schipper, B.C. (2015). Sex hormones and choice under risk. *Mimeo*. University of California, Davis.

Simonsohn, U., Nelson, L.D., Simmons, J.P. (2014). P-curve: a key to the file-drawer. *Journal of Experimental Psychology: General*, 143(2), 534-547.

Stenstrom, E., Saad, G., Nepomuceno, M.V., Mendenhall, Z. (2011). 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*, 51(4), 412–416.

Talarovičová, A., Kršková, L., Blažeková, J. (2009). Testosterone enhancement during pregnancy influences the 2D: 4D ratio and open field motor activity of rat siblings in adulthood. *Hormones and Behavior*, 55(1), 235-239.

Tester, N., Campbell, A. (2007). Sporting achievement: what is the contribution of digit ratio? *Journal of Personality* 75(4), 663-77.

Trivers, R., Manning, J.T., Jacobson, A. (2006). A longitudinal study of digit ratio (2D: 4D) and other finger ratios in Jamaican children. *Hormones and Behavior*, 49(2), 150-156.

Van Anders, S.M., Vernon, P.A., Wilbur, C.J. (2006). Finger-length ratios show evidence of prenatal hormone-transfer between opposite-sex twins. *Hormones and Behavior*, 49(3), 315–319.

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

Viscusi, W.K., Evans, W.N. (1990). Utility functions that depend on health status: estimates and economic implications. *The American Economic Review*, 80(3), 353-374.

Weber, E.U., Blais, A-R., Betz, N.E. (2002). A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*. 15(4), 263-290.

Zeckhauser, R.J., Viscusi, W.K. (1990). Risk within reason. *Science*. 248, 4955, 559-564.

Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., von Schoultz, B., Lindén A., Johannesson, M. (2009) A randomized trial of the effect of estrogen and testosterone on economic behavior. *Proceedings of the National Academy of Sciences,*106 (16) 6535-6538.

Zheng, Z., Cohn, M.J. (2011). Developmental basis of sexually dimorphic digit ratios. *Proceedings of the National Academy of Sciences,* 108, 16289–16294. doi: 10.1073/pnas.1108312108

Zhong, S., Chew, S.H., Set, E., Zhang, J., Xue, H., Sham, P.C., Ebstein, R.P., Israel, S. (2009). The heritability of attitude toward economic risk. *Twin Research and Human Genetics*, 12(1), 103-107.

Appendix A

Table A1. Pairwise correlations between the main variables, male subjects only.

	R2D:4D	**L2D:4D**	**RP**	**RA**
R2D:4D	1			
L2D:4D	0.630*** (0.000)	1		
RP	-0.126* (0.070)	-0.188*** (0.007)	1	
RA	-0.031 (0.639)	-0.039 (0.564)	0.288*** (0.000)	1

Note: * = p < 0.1, ** = p < 0.05, *** = p < 0.01.

Table A2. Pairwise correlations between the main variables, female subjects only.

	R2D:4D	**L2D:4D**	**RP**	**RA**
R2D:4D	1			
L2D:4D	0.748*** (0.000)	1		
RP	-0.098** (0.036)	-0.040 (0.389)	1	
RA	0.037 (0.419)	0.0171 (0.710)	0.148*** (0.001)	1

Note: * = p < 0.1, ** = p < 0.05, *** = p < 0.01.

Table A3: RP and individual characteristics: all subjects (OProbit).

RP	*m1*	*m2*	*m3*
Female	-0.211**		-0.215**
	(0.093)		(0.094)
Chinese		-0.210**	-0.198**
		(0.099)	(0.100)
S-asian		-0.000	0.019
		(0.124)	(0.124)
Black		-0.099	-0.133
		(0.216)	(0.222)
Other		0.120	0.115
		(0.141)	(0.140)
Observations	664	664	664
Pseudo R^2	0.0027	0.0040	0.0064

Note: Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors of the cut-points have been omitted.

Table A4. RA and individual characteristics: all subjects (OProbit)

RA	*m1*	*m2*	*m3*
Female	-0.213***		-0.191**
	(0.083)		(0.083)
Chinese		-0.287***	-0.276***
		(0.091)	(0.092)
S-Asian		-0.107	-0.089
		(0.115)	(0.115)
Black		0.201	0.172
		(0.210)	(0.207)
Other		-0.101	-0.109
		(0.135)	(0.135)
Observations	704	704	704
Pseudo R^2	0.0022	0.0043	0.0060

Note: Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors of the cut-points have been omitted.

Table A5. RA, L2D:4D and individual characteristics: all subjects (OProbit)

RA	*m1*	*m2*	*m3*	*m4*	*m5*
L2D:4D	-0.679	-1.141	-1.102	-0.901	-0.430
	(1.200)	(1.215)	(2.035)	(1.203)	(1.220)
Female		-0.211**	-1.668		-0.185**
		(0.084)	(2.455)		(0.085)
L2D:4D*Female			1.511		
			(2.549)		
Chinese				-0.293***	-0.279***
				(0.091)	(0.092)
S-Asian				-0.106	-0.089
				(0.115)	(0.115)
Black				0.194	0.170
				(0.213)	(0.209)
Other				-0.107	-0.112
				(0.135)	(0.135)
Observations	704	704	704	704	704
Pseudo R^2	0.0001	0.0022	0.0024	0.0045	0.0061

Note: Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors of the cut-points have been omitted.

Table A6: RP, R2D:4D and individual characteristics: all subjects, female subjects, male subjects (OProbit)

RP	*m1: all*	*m4: all*	*m1: female*	*m4: female*	*m1: male*	*m4: male*
R2D:4D	-3.939***	-4.457***	-3.081**	-3.544**	-4.230*	-4.461*
	(1.227)	(1.233)	(1.511)	(1.559)	(2.357)	(2.298)
Chinese		-0.248**		-0.250**		-0.211
		(0.101)		(0.122)		(0.184)
S-Asian		-0.001		-0.035		0.133
		(0.124)		(0.148)		(0.239)
Black		-0.175		0.081		-0.395
		(0.217)		(0.301)		(0.220)
Other		0.119		0.018		0.263
		(0.140)		(0.1721)		(0.241)
Observations	664	664	458	458	206	206
Pseudo R^2	0.0050	0.0102	0.0030	0.0072	0.0049	0.0146

Note: Standard errors in parentheses; * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors of the cut-points have been omitted.

Table A7: RP, L2D:4D and individual characteristics: all subjects, female subjects, male subjects (OProbit).

RP	*m1: all*	*m4: all*	*m1: female*	*m4: female*	*m1: male*	*m4 : male*
L2D:4D	-3.354***	-3.645***	-1.278	-1.579	-5.694***	-5.561***
	(1.215)	(1.220)	(1.541)	(1.574)	(2.065)	(2.046)
Chinese		-0.238**		-0.230*		-0.218
		(0.101)		(0.122)		(0.184)
S-asian		0.001		0.032		0.135
		(0.123)		(0.148)		(0.232)
Black		-0.145		0.123		-0.388
		(0.212)		(0.299)		(0.302)
Other		0.096		-0.001		0.192
		(0.141)		(0.171)		(0.247)
Observations	664	664	458	458	206	206
Pseudo R^2	0.0037	0.0083	0.0005	0.0042	0.0112	0.0195

Note: Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors of the cut-points have been omitted.

Table A8. RA, R2D:4D and individual characteristics: all subjects, female subjects, mal subjects (OProbit).

RA	*m1: all*	*m4: all*	*m1: female*	*m4: female*	*m1: male*	*m4: male*
R2D:4D	-0.370	-0.523	0.954	0.701	-0.892	-1.022
	(1.127)	(1.129)	(1.415)	(1.431)	(2.205)	(2.034)
Chinese		-0.291***		-0.304***		-0.222
		(0.091)		(0.110)		(0.167)
S-Asian		-0.106		-0.089		-0.131
		(0.115)		(0.140)		(0.205)
Black		0.193		-0.069		0.463
		(0.211)		(0.288)		(0.299)
Other		-0.101		-0.215		-0.053
		(0.135)		(0.171)		(0.218)
Observations	704	704	478	478	226	226
Pseudo R^2	0.0000	0.0043	0.0002	0.0041	0.0002	0.0067

Note: Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors of the cut-points have been omitted.

Table A9. RA, L2D:4D and individual characteristics: all subjects, female subjects, male subjects (OProbit)

RA	*m1: all*	*m4: all*	*m1: female*	*m4: female*	*m1: male*	*m4: male*
L2D:4D	-0.679	-0.901	0.422	0.138	-1.057	-1.349
	(1.200)	(1.203)	(1.538)	(1.546)	(2.050)	(2.041)
Chinese		-0.293***		-0.307***		-0.225
		(0.091)		(0.111)		(0.167)
S-Asian		-0.106		-0.089		-0.133
		(0.115)		(0.140)		(0.205)
Black		0.194		-0.079		0.474
		(0.213)		(0.289)		(0.302)
Other		-0.107		-0.211		0.035
		(0.135)		(0.170)		(0.222)
Observations	704	704	478	478	226	226
Pseudo R^2	0.0001	0.0045	0.0000	0.0040	0.0003	0.0070

Note: Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Estimated coefficients and standard errors of the cut-points have been omitted.

Table A10. Pairwise correlations between the main variables with Bonferroni (1935) correction.

	R2D:4D	L2D:4D	RP	RA
R2D:4D	1			
L2D:4D	0.719*** (0.000)			
RP	-0.126*** (0.007)	-0.108** (0.031)		
RA	-0.010 (1.000)	-0.021 (1.000)	0.204*** (0.000)	1

Note: Standard errors in parentheses. * = $p < 0.1$, ** = $p < 0.05$, *** = $p < 0.01$.

Appendix B

Subject consent form for Digit Ratio measurement

Please read this consent form carefully and ask as many questions as you like before you decide whether or not you want to participate in the next measurement. Before you leave the laboratory today, we are asking everyone to take a measure called the digit ratio. This ratio is calculated by combining the length of your 2nd and 4th finger, and it has been shown in various scientific studies to correlate with people's behaviour in the laboratory. The most efficient and reliable way of measuring the ratio is by scanning someone's hand on a flatbed scanner.

As with all responses during our experiments, we will collect your digit ratio completely anonymously. No-one, not even the researcher in charge of the study, will be able to link your digit ratio to your identity, name, and personal information. As such, we will not be able to share your digit ratio with anyone, including you.

There are no risks to you from this research and no foreseeable direct benefits. It is hoped that the research will benefit others (or science) who wish to understand behaviour and decisions. The researcher in charge of today's study has collected digit ratio data in the LSE Behavioural Research Lab before. The image data will only be used for calculating the digit ratios, and it will be stored on an encrypted hard drive with no access to any external networks, kept in a secure storage space which will only be accessible by the researchers directly involved in this project.

If you have any questions about anything, please ask them now and/or contact the researcher in charge of the study: [contact details provided]. If you agree to provide a digit ratio measure, please continue.

****

I have read and understand this consent form and I am willing to provide a digit ratio measure

________________________	____________________________	____________
Signature	Name (please print)	Date