Essays in behavioural economics

Buser, T.

tinbergen institute

Thomas Buser

Essays in Behavioural Economics

Universiteit van Amsterdam

Essays in Behavioural Economics Thomas Buser

538

Most of the economic literature takes preferences as given. Economists use them as the building blocks of their models, estimate them in lab experiments and correlate them with life outcomes. But we only rarely ask about their origins. How come that fundamental preferences such as risk aversion, reciprocity, inequality aversion and altruism vary so strongly across individuals? Are they biologically predetermined or can we shape them through upbringing and education?

This dissertation aims to take a step towards answering these questions. Chapter two investigates whether the willingness to compete in a tournament varies over the menstrual cycle and with the intake of hormonal contraceptives. Chapters three and four correlate choices in social preference games with physical markers for underlying neural differences. Chapter five deals with the consequences rather than the origins of individual differences in economic preferences and investigates whether gender differences in academic career choices can be explained by gender differences in the willingness to compete.

Thomas Buser holds a Bachelor degree from the University of Lausanne and a Master degree from the University of Warwick. He graduated from the Tinbergen Institute MPhil program in 2009 and conducted his PhD research at the University of Amsterdam.

Essays in Behavioural Economics

ISBN 978 90 361 0312 1

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 538 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Essays in Behavioural Economics

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 4 september 2012, te 10:00 uur

door

Thomas Buser

Geboren te Basel, Zwitserland

Promotiecommissie

Promotoren:	Prof. dr. Hessel Oosterbeek
	Prof. dr. Erik Plug
Overige leden:	Prof. dr. Uri Gneezy
	Dr. Guillaume Hollard
	Prof. dr. Henriette Maassen van den Brink
	Prof. dr. Theo Offerman
	Prof. dr. Frans van Winden

Faculteit Economie en Bedrijfskunde

Acknowledgements

The defence of my PhD thesis is a great occasion to express my gratitude towards the people who have made the last five years in Amsterdam such a great time, both privately and professionally. First of all I want to thank Hessel, not only for being an amazing supervisor (I listened to his advice more often than he might think), but for being a great friend. I hope we will be talking, drinking, travelling, cooking, playing cards and working together for a long time to come. Many thanks also go to Erik, whose advice and insight improved many a research design and paper, and who is always great fun to be around.

Apart from my supervisors, the person who contributed the most to this thesis (and my fledgling career as an economist) is Muriel. I would like to thank her for her repeated invitations to Stanford University, her hospitality, her advice and her confidence in my skills. I would also like to thank my coauthors on past and current projects which are not included in this thesis: Anna, Jose, Noemi, it is a pleasure to work with you and I have learned a lot from each of you.

I would also like to thank my other past and future colleagues at TIER – Adam, Diana, Jona, Monique, Nadine, Nienke and Nina – for their help and advice as well as fun times. I am excited to keep working with you in the future. I am also grateful to the people at CREED who were always helpful and supportive of my infringement on their experimental turf.

My time at TI and UvA has not only led to this thesis but also to some lasting friendships without which my years in Amsterdam would have been a much less memorable experience. Christel, Eva, Marcel, Marcel and Stephen: we shared innumerable great times (and, with some of you, a house) together and although you soon will be scattered over various continents, I know we will always stay in touch. Luckily for my own mental sanity (if probably not theirs) I have also found great friends who do not spend their days researching obscure economic problems. Thanks a lot to Alexandra, Amanda, Carol, Floris, Ingrid, Job, Martin, Val and all the other people whom I have shared food, drinks and conversations with during the past five years.

My years of studying which led up to this thesis would not have been possible without my parents. They were always supportive when I put off serious employment to pursue yet another academic degree. For this I will always be grateful.

But most of all I want to thank my wonderful wife Lygia who dropped her life in Brazil to come live with me in this cold and unknown country called the Netherlands. I cannot wait to find out what our future together holds in store for us.

Contents

Chapter 1

Introduction

If this thesis has a unifying theme, it is that all chapters are essentially about individual differences in economic preferences. Most of the economic literature takes preferences as given. Economists use them as the building blocks of their models, estimate them in lab experiments and correlate them with life outcomes. But we only rarely ask about their origins. How come that fundamental preferences such as risk aversion, reciprocity, inequality aversion and altruism vary so strongly across individuals? Are they biologically predetermined or can we shape them through upbringing and education?

This question is an important one. As Chapter 5 of this thesis illustrates, economic preferences influence some of the most important life decisions, including the choice of study major and career, and it is therefore important to know how they are determined. This is the premise of Chapters 2 to 4. In these chapters, we correlate physical markers for hormonal and neurological differences with choices in laboratory experiments. A significant correlation, we argue, is evidence that nature plays a role in shaping our economic preferences.

In Chapter 2, we investigate the impact of hormonal fluctuations on competitiveness. Previous lab studies have found that when given a choice, men are generally attracted to competition while women shy away from it.[1] One can easily extrapolate this to choices in the labour market and it becomes obvious that the gender difference in the taste for competition could be an explanation for gender differences in career choices. But why is it the case that men are more attracted to competitive environments than women? Is it the different upbringing that girls receive from parents and educational institutions or is there an innate difference in competitiveness? One important biological difference between men and women consists in the differential exposure to male and female sex hormones. Testosterone levels are much higher in men while women are exposed

[1] See Niederle and Vesterlund (2011) for an overview.

to much higher levels of oestrogen and progesterone. Can these hormonal differences partially explain the gender difference in competitiveness?

We take a step towards answering this question by taking advantage of a natural experiment in the exposure to female sex hormones which is going on all around us, namely the menstrual cycle. Women who do not take hormonal contraceptives are subject to strong and predictable fluctuations in progesterone and oestrogen (plus a number of other hormones). Women who take hormonal contraceptives are subject to exogenous variations in sex hormone levels as well, as the levels of oestrogen and progesterone decline dramatically during the seven-day pill-break. Using a simple and well-established design where subjects choose between piece-rate and tournament pay for a simple arithmetic task[2], we show that hormone levels strongly predict the likelihood of making the competitive choice of entering the tournament. Women who take the pill are roughly twice as competitive during the pill-break, when hormone levels are low, compared to the pill intake phase, when hormone levels are high. Women who experience a natural cycle are roughly half as competitive during the days following ovulation than during the rest of the cycle. Together, these effects hint at a negative impact of the hormone progesterone on competitiveness.

Social preferences such as trust, altruism and reciprocity also vary strongly between individuals. Again one can ask to which extent these differences are biologically determined. In Chapters 3 and 4, we let subjects participate in a range of social preference games commonly used in the economic literature (the trust game, the ultimatum game, the public good game and the dictator game) and correlate their choices with physical markers for underlying neurological and hormonal differences. We argue that a significant correlation would indicate that there is a partial biological basis for social preferences.

In Chapter 3, we concentrate on the effects of prenatal hormone exposure. The extent to which we are exposed to the hormones testosterone and oestrogen in utero varies strongly between individuals within each gender. The strength of exposure is thought to have permanent effects on the structure of the brain, and consequently behaviour and preferences later in life. Stronger prenatal exposure to testosterone is, for example, thought to lead to a brain which is worse at empathising but better at dealing with systems.[3] Curiously, the relative strength of testosterone to oestrogen exposure also determines the relative length of the index and ring fingers, which is fixed before birth.[4] By simply asking our subjects whether they have a longer index or ring finger, we can therefore obtain a crude proxy for the neurological differences associated with prenatal hormone exposure. We find that subjects with finger ratios indicating a stronger exposure to testosterone are less generous in all games. We speculate that a negative effect of prenatal testosterone exposure

[2]See Niederle and Vesterlund (2007).
[3]See Baron-Cohen et al. (2004) for an overview.
[4]See Manning (2002) for an overview.

on empathy lies behind these results. We also ask the female subjects a range of menstrual cycle questions and find that, like competitiveness, social preferences vary over the menstrual cycle.

In Chapter 4, we correlate the same social preference data with handedness. Left-handers differ neurologically from right-handers (and these differences vary between men and women).[5] We find that left-handed men are significantly more generous when recipients can reciprocate and exhibit stronger positive reciprocity themselves. Left-handed women, on the other hand, are significantly less altruistic. Information on handedness is available in surveys and we can therefore check whether our results carry over to behaviour outside of the lab. Using Dutch and American survey data, we find significant effects of handedness on self-rated trust and altruistic behaviour including donating money and doing volunteer work. Combined with the results presented in Chapter 3, this is evidence that there is a neural basis for social preferences.

Chapter 5 deals with the consequences rather than the origins of individual differences in economic preferences. Here, we delve deeper into the subject of gender differences in competitiveness. The initial finding of a gender gap in competitiveness[6] has spawned a sizeable and still growing literature based on the extrapolation of the lab findings to labour market settings. If women dislike being in competitive environments, this could cause them to pick different careers and could ultimately be an explanation for labour market differences between men and women. We test the validity of this extrapolation by correlating an experimental measure of competitiveness with the first important career choice of secondary school students in the Netherlands. At the age of 15, these students have to pick one out of four study profiles: a science-oriented profile, a health-oriented profile, a social science-oriented profile and a humanities-oriented profile. These profiles are a strong predictor of the probability of going to university and of the chosen university major. Choices of boys and girls show clear differences. Boys concentrate in the science-oriented profile, which is widely regarded as the most prestigious and challenging profile. Girls concentrate in the less prestigious health- and humanities-oriented profiles. We replicate the finding that boys are much more competitive than girls. We also find that competitiveness significantly affects profile choice. Gender differences in competitiveness can account for around 20 percent of gender differences in career choices. This supports the extrapolation of laboratory findings on competitiveness to labor market settings and suggests that differences in competitiveness are indeed partially responsible for the different career paths of men and women.

[5] See McManus (2002) for an overview.

[6] See Niederle and Vesterlund (2007).

Chapter 2

The Impact of the Menstrual Cycle and Hormonal Contraceptives on Competitiveness[1]

2.1 Introduction

Selection procedures for high paying jobs, promotions, and wage increases are often based on tournament-like competition which is believed to select the highest performers for the task at hand. A growing literature, however, demonstrates that there are individual differences in competitiveness that determine selection in and out of tournaments independently from performance. Some individuals simply seem to dislike being in competitive situations. The strongest evidence comes from the experimental literature on gender and competitiveness which finds that women tend to dislike competition while men actively seek it. The aim of this study is to determine whether the menstrual cycle and intake of hormonal contraceptives have an impact on the competitiveness of women. As both lead to predictable hormonal fluctuations, such an impact would be an indication that individual differences in competitiveness are at least partially caused by biological factors.

Most experimental studies on competitiveness have subjects perform a simple task whereby the compensation scheme is varied between a non-competitive piece rate and a competitive tournament scheme. Niederle and Vesterlund (2007) find that, when given a choice, 73 percent of men but only 35 percent of women opt to compete. Gneezy et al. (2003) moreover find that men significantly increase effort when the compensation scheme for a task becomes more competitive while

[1] The research in this chapter has been published in Buser (2011).

women show no reaction.[2] There is evidence that nurture can explain at least part of these gender differences. Gneezy et al. (2009) conduct the same compensation choice experiment with subjects from a patriarchal society (the Maasai of Tanzania) and subjects from a matrilineal society (the Khasi of India). While the Maasai exhibit the same gender gap in competitiveness found in Western societies, the roles are reversed in the Khasi sessions, though the authors explicitly mention the possibility that nature, as well as nurture, may play a role in this reversal. Further evidence comes from Cardenas et al. (2011) who find that gender differences in competitiveness vary across countries and may be correlated with gender stereotypes.

How much of a role nature plays in determining attitudes towards competition is still largely an open question. In a rare study on the impact of hormones, Apicella et al. (2011) find no effect of testosterone on tournament entry in men. But in other areas of economic behaviour, testosterone has been associated with lower offers and more rejections in the ultimatum game (Burnham, 2007; Zak et al., 2009), increased financial risk taking (Apicella et al., 2008), and the likelihood for MBA students to seek out a career in finance (Sapienza et al., 2009). Also, the hormone oxytocin increases giving in the trust game (Kosfeld et al., 2005) and the ultimatum game (Zak et al., 2007).[3] Zethraeus et al. (2009), on the other hand, find no impact of testosterone and oestrogen levels in a range of games measuring altruism, trust, fairness, and risk aversion.

The impact of the menstrual cycle on economic decision making has so far only been analysed in the context of sealed bid first-price auctions. Chen et al. (2009) find that bidding fluctuates over the menstrual cycle for users of hormonal contraceptives only. In a replication, Pearson and Schipper (2011b) also find significant, but partially contradictory, fluctuations. Since the first version of this study has been released, one other study concerned with the impact of the menstrual cycle on competitiveness has appeared (Wozniak et al., 2010). We will provide a more detailed discussion of these studies and a comparison of results in Section 2.5.

We hypothesise that competitiveness is related to fluctuations in female sex hormones and that it consequently fluctuates over the menstrual cycle and with contraceptive intake. Moreover, we expect competitiveness to fall when sex hormone levels are high and to rise when they are low. Such a finding would indicate that innate differences can explain a significant part of the gender gap in competitiveness. If the divergence between the competitive behaviour of men and women is due solely to nurture, on the other hand, we would expect to observe no effects.

Our results strongly confirm our hypotheses. Making use of the diverging patterns of oestrogen and progesterone secretion over the menstrual cycle, we also find that the fluctuations in competitiveness are most strongly correlated with fluctuations in progesterone levels. We consider three

[2] See Croson and Gneezy (2009) for a review of gender differences in lab and field experiments covering the areas of risk aversion, competitiveness, and social preferences.

[3] Fehr (2009) reviews further evidence of biological and other factors influencing trusting behaviour.

possible indirect pathways for the effect of the menstrual cycle and contraceptives on competitiveness: via an impact on risk aversion, via an impact on maths performance, and via an impact on overconfidence. None of these hold up to the data.

The next section of this chapter describes which variables we use to capture the relevant features of the menstrual cycle and of hormonal contraceptives. Section 2.3 provides further details about the experimental design, and Section 2.4 describes the sample. Section 2.5 presents the main results and Section 2.6 reports the findings regarding possible pathways. Section 2.7 concludes.

2.2 Measurement of menstrual cycle phases and hormonal contraceptives

The medical literature commonly divides the menstrual cycle into five phases across which the levels of female sex hormones fluctuate according to a predictable pattern (see e.g. Richardson, 1992).[4] These phases and the fluctuations of oestrogen and progesterone assuming a regular 28-day menstrual cycle are illustrated in Figure 2.1. Women using hormonal contraceptives are subject to a different 28-day cycle wherein a 21-day intake period, which is characterised by constant daily doses of an artificial oestrogen and an artificial progestin[5], is followed by a 7-day break. Oestrogen excretion by the body is markedly reduced in women taking hormonal contraceptives and progesterone excretion ceases almost completely (Rivera et al., 1999). This leads to a regular pattern whereby hormone levels are high during the 21-day intake period and low during the 7-day pill break.

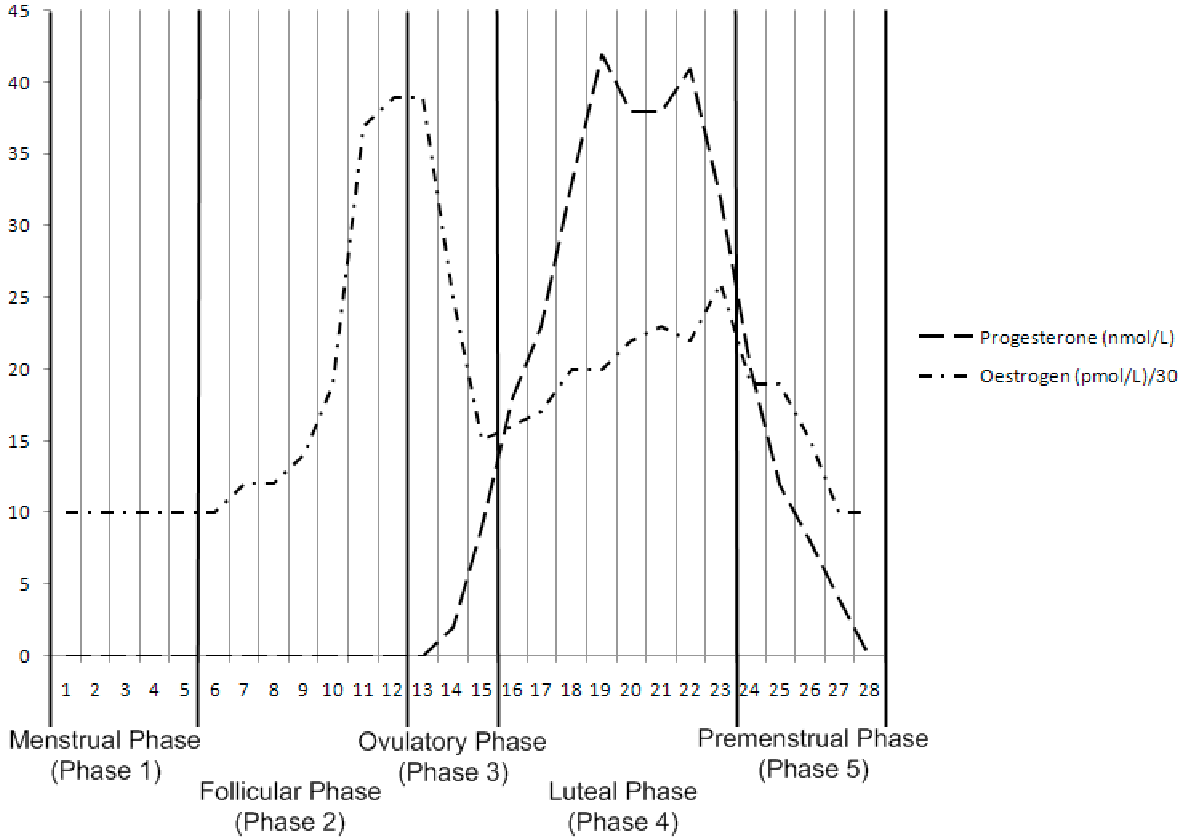
We elicit the expected beginning of the next menstruation and use this to allocate subjects experiencing a natural cycle to one of the five cycle phases. Cycle length varies across individuals whereby the follicular phase is the most variable while the length of the ovulatory, luteal, and premenstrual phases is relatively fixed (Hampson and Young, 2008). Allocation between phases 2 to 4 should therefore be less affected by varying cycle lengths. Moreover, we ask subjects whether they are currently experiencing menstrual bleeding and use this to allocate subjects between phases 1 and 2. We also construct two continuous variables representing the expected oestrogen and progesterone levels given the day of the cycle a subject is currently in.[6] As the pill break coincides with the menstrual period for hormonal contraceptive takers, we define as pill break subjects those who are 20 or more days away from their next menstruation. This means that subjects are counted

[4] Levels of testosterone are virtually constant over the cycle.

[5] A progestin is a synthetic hormone that has effects similar to progesterone.

[6] The average daily plasma hormone levels over the menstrual cycle are obtained from Chabbert Buffet et al. (1998) and are illustrated in Figure 2.1. We did not take any direct hormone measurements.

Figure 2.1: Hormone Levels over the Menstrual Cycle



Hormone levels are obtained from Chabbert Buffet et al. (1998); oestrogen levels have been reduced by a factor of thirty.

as high-hormone from the day after they take the first pill of a new package and as on the break from the day after they take the last.

For our analyses using the whole pooled sample, we divide subjects into high-oestrogen and low-oestrogen, as well as high-progesterone and low-progesterone individuals. The high oestrogen phase corresponds to cycle phases two and four while the high progesterone phase coincides with the fourth phase. For subjects taking hormonal contraceptives, the high-oestrogen and high-progesterone phases are congruent and coincide with the pill-intake phase.

Using self-reported menstrual cycle data introduces measurement error. Women often misestimate their cycle length (Small et al., 2007) and cycle length tends to vary around the mean over time (Creinin et al., 2004). Moreover, while the follicular phase varies most with the length of the cycle, there is some variability in the length of other phases too (Stern and McClintock, 1998). As women estimate their cycle length correctly on average (Creinin et al., 2004), this leads to classical measurement error, introducing random noise which biases any of the estimated effects towards zero. Measurement error is less of a problem for our estimates based on the sample of contraceptive takers for whom a regular 28-day cycle is virtually guaranteed and whom are divided into only two groups. Also, the core of our analysis is not based on the five phases but divides the whole sample into a high and a low hormone group. The robustness checks reported in Section 2.5 confirm that these estimates are unlikely to be strongly affected by misallocations. Conversely, we expect that the division into five cycle phases and the constructed continuous hormone variables are the measures most affected by measurement error as an error of only a few days can change expected levels a lot. But it is also important to stress that we observe the same patterns irrespective of which indicator – cycle phase dummies, contraceptives, or daily levels and changes – we use.

2.3 Experimental design

The design of the competition part closely follows the methodology of Niederle and Vesterlund (2007). Subjects are divided into groups of four and are given 5 minutes to solve as many sums of five two-digit numbers as they can, a task for which no gender differences have been observed. In a first round, subjects receive a piece rate of 1€ for each correct answer. In a second round, they compete in a tournament where the highest performer of each group receives 4€ per correct answer while the rest receive nothing. Being informed about her absolute but not her relative performance, each subject then decides which of the two compensation schemes she wishes to apply in a third round. Subjects going for the tournament in round three receive 4€ per correct answer if they score higher than the best of their group mates did in round two. A random pick of one of the rounds is relevant for payment. Finally, we elicit subjects' beliefs concerning their group rank for each task, paying 2€ for each correct guess.

To measure attitudes towards risk, we conducted a simple lottery choice experiment which follows the methodology of Eckel and Grossman (2002). Subjects make a single choice between a sure payoff of 8 Euros and four 50/50 lotteries with linearly increasing riskiness and expected payoffs: 12/6, 16/4, 20/2, 24/0. The choice of lottery then serves as an indicator of the risk aversion of the subject, yielding a discrete variable ranging from 1 (sure thing) to 5 (highest expected payoff/highest risk option).[7]

Our study is conducted on an all-female sample. Given that women compete against both sexes in the labour market, a mixed sample may seem more natural. On the other hand, there is a large literature showing that women react differently to men at different points of their cycle for reasons not exclusively linked to competitiveness. Women are, for example, more attracted to masculine features during the fertile part of the menstrual cycle (Penton-Voak and Perrett, 2000) and less when progesterone levels are high (Jones et al., 2005). Reactions to male body odour (Thornhill and Gangestad, 1999) and the likelihood of extra-pair copulations (Bellis and Baker, 1990) also vary over the cycle. The presence of male subjects would consequently introduce a confounding factor, making it less clear whether fluctuations in behaviour over the cycle can be interpreted as changes in underlying competitiveness, the main interest of this study.

Menstrual cycle details are sensitive information to ask and a female assistant was therefore present at all sessions and was responsible for all interactions with the subjects concerning the post-experimental questionnaire. In the end, selective non-response turned out not to be a problem as all subjects chose to answer the questions. Subjects were paid for one randomly selected part and received a fixed fee of €10.[8] Seven sessions were conducted at the CREED computer lab at the University of Amsterdam in June 2009 and the experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

[7]We also measured risk attitudes through the methodology designed by Holt and Laury (2002). The two risk measures are highly correlated. But since the Holt-Laury measure is more complicated for subjects to grasp – leading some subjects to make inconsistent choices – we only use the results obtained with the Eckel-Grossman methodology in this study. Using the Holt-Laury measure instead or eliminating the subjects who made inconsistent decisions does not change any of our conclusions concerning risk attitudes.

[8]Apart from the competitiveness and risk parts, subjects also participated in a social preference and public goods part. The social preference games were all one-shot and sample sizes – with only half of the subjects being a giver and half a receiver – were therefore too small to be useful. The results of a follow-up study concentrating on social preferences are contained in the following two chapters.

Table 2.1: Descriptive Statistics

	Sample	Natural cycle	Pill takers
Age	23.2	24.0	22.5
Economics	42.1%	50.0%	34.0%
Female	100%	100%	100%
Nationality			
Dutch	47.7%	29.6%	66.0%
Other European	43.0%	57.4%	28.3%
Latin American	3.7%	3.7%	3.8%
Other	5.6%	9.3%	1.9%
N	107	54	53

The binary indicator Economics is equal to one for subjects majoring in economics, econometrics or finance.

2.4 Data

The sample consists of 120 female university students of which we have to drop 13 who state not to experience a menstrual cycle at all.[9] Of the remaining 107 subjects, 53 use hormonal contraceptives. Table 2.1 shows descriptive statistics including age, nationality, and study major. It is apparent that contraceptive takers and non-takers differ along most dimensions. This does not affect our results as we only compare high and low hormone subjects within each group.

Table 2.2 contains the actual and expected distribution of subjects across menstrual cycle phases and between the pill-intake and pill-break phases. Selective attrition due to menstruating subjects staying away is not a significant problem: a χ^2-test cannot reject equality of the observed distribution and the theoretical distribution (p=0.50). Only subjects in the premenstrual phase, in which premenstrual symptoms such as cramps can occur, are underrepresented. But this does not affect our conclusions as our regression results are robust to the exclusion of phase five subjects. There is no attrition problem for subjects using hormonal contraceptives: the number of subjects on the pill-break is exactly equal to the expected number.[10] We cannot reject the null hypothesis of subjects being randomly distributed across the different phases of the cycle with respect to their age (Kruskal-Wallis test: p=0.46) and nationality (Fisher's exact test: p=0.48). The same is true for users of hormonal contraceptives when it comes to assignment to the pill break (p=0.95 and p=1.00 respectively).

[9]Subjects gave a range of reasons for not experiencing a menstrual cycle including using intra-uterine devices (which completely suppress menstruation and make menstrual cycle assessment impossible), transsexuality and pregnancy.

[10]Neither do subjects on the pill break differ from subjects in the pill-intake phase in the characteristics of the contraceptives they take: Fisher's exact test returns a p-value of 0.99 with respect to progestin type and the Wilcoxon rank-sum test returns a p-value of 0.90 with respect to oestrogen dosage.

Table 2.2: Subjects by Menstrual Cycle Phase

Menstrual Cycle or Pill Cycle Phase	Number of Subjects	Expected Number of Subjects
Menstrual Phase (5 days)	11	10
Follicular Phase (7 days)	15	13
Peri-Ovulatory Phase (3 days)	9	6
Luteal Phase (8 days)	15	15
Premenstrual Phase (5 days)	4	10
Pill Break (7 days)	13	13
Pill Intake Phase (21 days)	40	40

2.5 Results

The proportion of subjects opting for competition in round three is 44.9 percent. If our hypotheses are correct we can expect competitiveness to vary across the five menstrual cycle phases. Also, we can expect competitiveness to be lower during the luteal phase when both progesterone and oestrogen levels are particularly high. Competitiveness should also be lower for subjects currently on the pill.

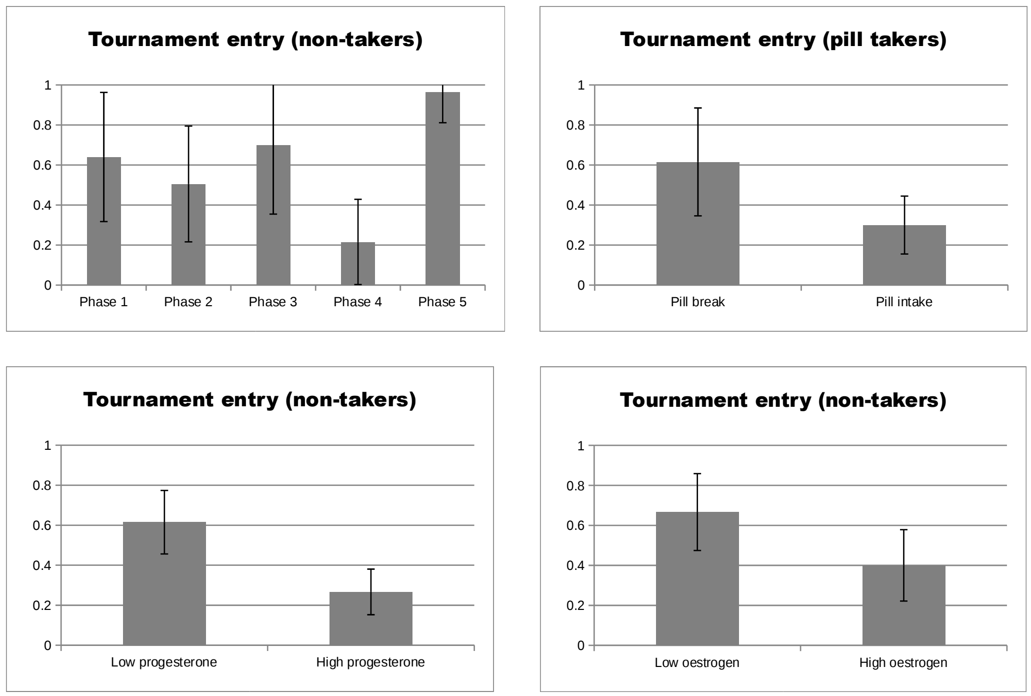
Tournament entry rates indeed vary significantly across cycle phases (ANCOVA with demographic controls[11]: p=0.03). Competitiveness is particularly low for subjects in the high-progesterone luteal phase (Wilcoxon rank-sum test: p=0.02) and is also lower in phases two and four combined, which represents the high-oestrogen period (p=0.05). Effects are equally strong for the sample of contraceptive users, who are significantly less competitive during the pill-intake phase (p=0.04).[12] Both samples approximately exhibit a doubling in the entry rate between the high to the low hormone phase. These results are illustrated in Figure 2.2.

Figure 2.3 and Table 2.3 show regression results for differences in competitiveness between the high and low hormone phases for the whole sample including pill takers and non-takers. We can see that tournament entry is about twice as high during the low progesterone phase than during the high progesterone phase with the difference between the low and high oestrogen phases being similarly large. These effects are robust to the inclusion of controls and significant at the 0.01-level throughout. As a robustness check, we ran the same regressions changing the lower and upper boundaries of the high progesterone, high oestrogen and pill-intake phases by up to three days in either direction. The hormone effects always stay significant and we conclude that the estimates are robust to cycle phase misallocations. The oestrogen results are also robust to the

[11] Demographic controls in all our analyses consist of age, educational background (as defined in Table 2.1), and nationality.

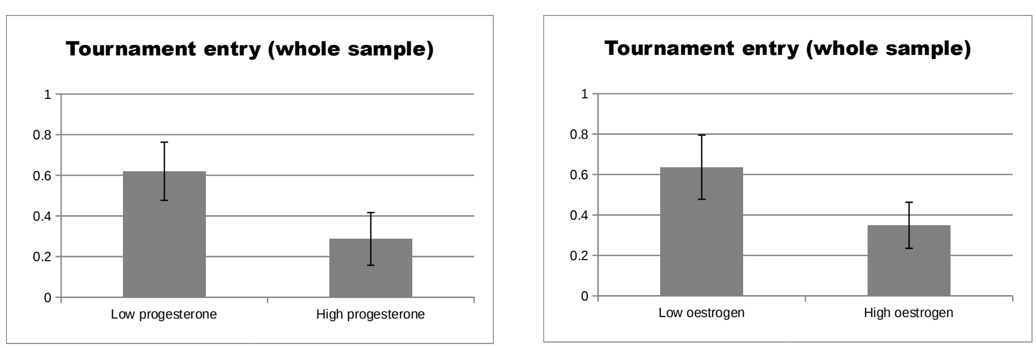
[12] The significance of this difference is confirmed by an ANCOVA model with demographic controls (p=0.04).

Figure 2.2: Hormones and Tournament Entry Rates for Pill Takers and Non-Takers (with 95%-confidence intervals)



All coefficients and 95%-confidence intervals are obtained from OLS regressions with robust standard errors. The phase-dummy regression controls for age, educational background, and nationality.

Figure 2.3: Hormones and Tournament Entry Rates for the Pooled Sample (with 95%-confidence intervals)



All coefficients and 95%-confidence intervals are obtained from OLS regressions with robust standard errors. The regressions control for a contraceptive taker dummy.

Table 2.3: Competitiveness Differences between High Hormone Subjects and Low Hormone Subjects

	(1)	(2)	(3)	(4)	(5)	(6)
Competitiveness						
High progesterone	-0.333***	-0.382***	-0.350***			
	(0.104)	(0.101)	(0.100)			
High oestrogen				-0.287***	-0.332***	-0.295***
				(0.101)	(0.099)	(0.099)
Contraceptive taker	0.018	0.100	0.135	-0.084	-0.007	0.036
	(0.104)	(0.099)	(0.091)	(0.096)	(0.097)	(0.091)
Risk aversion			0.087**			0.088**
			(0.036)			(0.036)
Performance			0.002			0.005
			(0.011)			(0.012)
Confidence			0.195**			0.182*
			(0.093)			(0.096)
Demographic Controls	no	yes	yes	no	yes	yes
Observations	107	107	107	107	107	107
R-squared	0.107	0.230	0.299	0.092	0.218	0.285

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; demographic controls consist of age, nationality, and study background; risk aversion is measured by the Eckel-Grossman scale, performance by the average performance of subjects in rounds one and two, and confidence by the belief to have been amongst the top two in one's group.

inclusion of phase 3 subjects in the high oestrogen group.[13]

So far, our results do not enable us to distinguish whether the fluctuations in competitiveness correlate more strongly with oestrogen levels or progesterone levels. Table 2.4 shows the results for linear probability models regressing tournament entry on daily expected oestrogen and progesterone levels and day-to-day changes for the sample of subjects experiencing a natural cycle. Columns (1) to (3) show that progesterone levels have a significant and negative impact while oestrogen levels are never significant. From Columns (4) to (6) we can see that day-to-day changes in progesterone levels are highly significant and negatively correlated with tournament entry while changes in oe-

[13]It is interesting to note that we find no impact of the menstrual cycle and contraceptives on the difference in arithmetic scores between round one (piece-rate) and round two (tournament) which, apart from learning effects, also incorporates the reaction of performance to the increase in the competitiveness of the compensation scheme. This is consistent with the finding of Niederle and Vesterlund (2007) that there is no gender gap in score improvement between rounds one and two.

strogen levels are marginally significant and negative as well. This means that competitiveness is lower when hormone levels are increasing and vice versa which explains the low levels of competitiveness during the luteal phase when progesterone is rising steeply and the high levels during the premenstrual phase when it decreases rapidly. This result is consistent with recent findings in endocrinology suggesting that changes in hormone concentrations might matter as much or more than levels in triggering hormone-induced processes.[14] The regression in Column (7) includes both levels and changes and confirms that competitiveness moves in step with progesterone (Wald test for joint significance of levels and changes: $p \ll 0.01$) rather than oestrogen (p=0.28).[15]

These results fit well with the wide variety of behavioural fluctuations over the menstrual cycle and the behavioural effects of progesterone documented in the medical literature.[16] They are also consistent with an evolutionary explanation according to which competitiveness is less desirable during the infertile phase of the menstrual cycle and during pregnancy (when hormone levels are high) than during the fertile phase (when competition for genetically well-endowed males is most important and hormone levels are low). The rush in progesterone occurring during the luteal phase signals the end of the fertile part of the menstrual cycle during which women are more likely to engage in extra-pair copulations (Bellis and Baker, 1990) and are more attracted to testosterone-related masculine facial features (Penton-Voak and Perrett, 2000). Jones et al. (2005) similarly show that women's commitment to their romantic relationship and attraction to femininity in male faces are positively and significantly correlated with progesterone levels.

Our findings are seemingly at odds with Wozniak et al. (2010) who report that women are more competitive during high hormone phases. But their design is very different as their subjects can choose between three options, adding a group scheme in which proceeds are shared equally. They find that participation in the group scheme is higher when hormones are low whereas tournament entry is higher during the high hormone phase. Piece-rate participation is actually slightly, if insignificantly, higher during the high phase too, which is in accordance with our results. It is impossible to know whether the subjects choosing the group scheme would have chosen the piece-rate or the tournament in our design and it is therefore difficult, if not impossible, to compare results. It is also unclear whether the group scheme is uncompetitive as subjects may feel compelled to live up to the expectations of the other group members or feel competitive pressure to perform better than them. Furthermore, entry into the group scheme might depend on factors such as altruism or a wish

[14] Kol and Homburg (2008), for example, propose that "changes in hormone concentrations carry significant biological messages, much more than a given level at a given time point".

[15] The results are robust to excluding premenstrual phase subjects. Given that premenstrual symptoms are hormone-driven, the under-representation of these subjects may cause bias. However, our main results carry through and our findings are therefore not an artifact of selective attrition of subjects in the fifth phase.

[16] de Wit et al. (2001), for example, find that exogenous administration of progesterone leads to feelings of sluggishness and a decrease in vigor.

Table 2.4: Natural Hormone Fluctuations and Competitiveness

Competitiveness	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Oestrogen (level)	0.004	0.005	0.004				0.003
	(0.009)	(0.009)	(0.010)				(0.008)
Progesterone (level)	-0.009*	-0.011***	-0.011**				-0.008**
	(0.005)	(0.004)	(0.005)				(0.004)
Oestrogen (change)				-0.015	-0.018*	-0.018*	-0.017
				(0.010)	(0.010)	(0.010)	(0.011)
Progesterone (change)				-0.034***	-0.042***	-0.042***	-0.036***
				(0.012)	(0.008)	(0.008)	(0.008)
Risk aversion			0.037			0.025	
			(0.059)			(0.053)	
Performance			0.005			0.016	
			(0.017)			(0.016)	
Confidence			0.039			-0.0570	
			(0.194)			(0.182)	
Demographic controls	no	yes	yes	no	yes	yes	yes
Observations	54	54	54	54	54	54	54
R-squared	0.054	0.209	0.220	0.111	0.283	0.303	0.319

Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$; demographic controls consist of age, nationality, and study background; risk aversion is measured by the Eckel-Grossman scale, performance by the average performance of subjects in rounds one and two, and confidence by the belief to have been amongst the top two in one's group.

to freeride which are unrelated to competitiveness and which may themselves be correlated with hormones. A further important design difference is that Wozniak et al. (2010) use a mixed gender sample. The remaining two experimental economics studies on the menstrual cycle look at bidding in first price auctions. Chen et al. (2009) find that contraceptive takers bid higher during and immediately after the pill break while there is no significant variation in bidding over the natural cycle. Pearson and Schipper (2011b) find that bidding is higher during the pre-menstrual and menstrual phases. Although both papers devote substantial space to discussing the differences in their results, it is worth pointing out that both find bidding to be higher during phases when hormone levels are low and which according to our results coincide with increased competitiveness.

2.6 Possible pathways

Our results show that the menstrual cycle and hormonal contraceptives have a significant impact on competitiveness. We will now investigate whether this effect is mediated by an impact on one of several possible determinants of competitiveness. We consider three possible pathways: via an impact on risk aversion, via an impact on mathematical abilities, and via an impact on overconfidence. None of these hypotheses hold up to the data.

2.6.1 Risk aversion

Datta Gupta et al. (2011) show that risk attitudes help determine women's competitiveness and there is a host of studies, including Eckel and Grossman (2002) and Powell and Ansic (1997), showing that women are more risk averse than men.[17] The regressions in Columns (3) and (6) of Table 2.3 show that an increase of one (on a five-point scale) in our risk taking indicator leads to an increase in the likelihood of competing of around nine percentage points. But neither the menstrual cycle phases nor hormonal contraceptives have a significant impact on risk aversion (ANCOVA with demographic controls; p=0.88 and p=0.18 respectively). However, pill takers are around one half of a standard deviation more risk averse during the intake phase and a doubling of the sample to n=106 would be enough to yield significance at the 5%-level. Future research on the effects of hormonal contraceptives on risk aversion seems therefore warranted. Chavanne and Gallup Jr (1998) and Bröder and Hohmann (2003) both study the impact of the menstrual cycle on risky behaviour and find a mid-cycle decrease. But these papers specifically look at behaviours which increase the risk of falling victim to rape and not at general risk attitudes. Also, risk aversion

[17] See Croson and Gneezy (2009) for a full survey of studies investigating gender differences in risk attitudes. The vast majority of surveyed papers find either that women are more risk averse than men or find no significant difference.

theoretically leads to higher bidding in first-price auctions and the findings of Chen et al. (2009) and Pearson and Schipper (2011b) are thus consistent with a negative impact of hormones on risk aversion. But these papers do not investigate directly whether the impact of the cycle on bidding is mediated via an impact on risk attitudes.

2.6.2 Mathematical ability

Average performance is 9.6 correct answers in round one and 11.5 in round two ($p<0.01$; one-sided t-test). We observe a further significant increase from the second to the third round even for those subjects choosing the piece rate ($p<0.01$), and it seems therefore more likely that the performance increase is due to learning effects than to a competition effect. The psychological literature has found some cognitive functions to vary over the menstrual cycle (Hampson and Kimura, 1992) and one could thus imagine that the same is true for the ability to solve sums. Diminished mathematical ability could obviously have a negative impact on subjects' readiness to compete. But a one-way ANCOVA model with demographic controls indicates that average mathematical performance shows no significant variation across the menstrual cycle phases ($p=0.75$) or between the pill-break and the pill-intake phase (p=0.21).[18] Moreover, absolute performance in rounds one and two, which is all the information subjects have at the moment of making their decision, has no impact on competitiveness. The regressions in Columns (3) and (6) of Table 2.3 show that the effect of the mean score from rounds one and two on the likelihood of competing in round three is both insignificant and negligibly small.[19]

2.6.3 Overconfidence

Niederle and Vesterlund (2007) find that confidence plays a significant but limited role in explaining whether an individual chooses to compete and that men are significantly more overconfident than women. Subjects are clearly overconfident: 67 percent believe to be amongst the top half of their group in round two and 41 percent of subjects overestimate their rank while only 21 percent underestimate it. We find some weak evidence that (over)confidence increases tournament entry. Subjects who overestimate their performance are 13 percent more likely to compete (one-sided

[18] Pill-takers do perform half a standard deviation worse during the pill intake phase though. A doubling of the sample is enough to yield significance at the 5%-level and an impact of hormonal contraceptives on performance can therefore not be rejected with confidence. We can still exclude an impact of contraceptives on maths scores as a possible pathway as performance has no effect on the likelihood of entering the tournament.

[19] Regressions without additional controls or with other measures of performance yield the same result. Scores from round two only and group ranks in rounds one and two are not significant in any specification when used to replace average performance. The same is true for dummies indicating an individual was the best or amongst the two best of her group.

t-test; p=0.10), but subjects who believe to be first – and who should therefore want to compete – are no more likely to enter the tournament than the rest (p=0.24). Subjects who believe to be amongst the two best in their group, however, are 16 percent more likely to compete (p=0.06). But conditional on performance neither the menstrual cycle phases (ANCOVA with controls; p=0.89) nor contraceptive intake (p=0.82) significantly affect the belief of subjects to be amongst the two best in their group.[20]

2.7 Conclusions

The labour market decisions of men and women are strikingly different, especially when it comes to the competitiveness of the chosen work environment. Simply put, men seem to actively seek competition while women tend to avoid it – a fact that is corroborated by several controlled experiments in the lab. Next to other explanations such as gender discrimination and conflicts between work and family life, this difference is likely one of the causes of the low number of women in top positions and the gender gap in wages.[21] It is therefore an important question whether these differences are purely a consequence of upbringing and education or whether biological differences between women and men play a role as well. Which policies we should adopt if we wish to tackle the gender imbalances in the labour market crucially depends on whether nature or nurture is at play.

Our results indicate that next to the cultural factors identified by Gneezy et al. (2009) amongst others, biological factors play a role in explaining gender differences in competitiveness. Multiplying the estimated coefficients for menstrual cycle phases two to five with their average duration, we find that women are 10.5 percentage points less likely to enter the tournament compared to a fictitious situation in which sex hormones are always at the low levels observed during the menstrual phase. And our regressions using daily expected hormone levels and changes indicate that the probability of entering the tournament is approximately fifty percent lower around day twenty of the menstrual cycle than during the menstrual phase. These back-of-the-envelope calculations indicate that the effect of hormones can account for a substantial part of the gender gap in competitiveness estimated by Niederle and Vesterlund (2007).

An interesting direction for future research could be to directly measure hormones by taking blood

[20]Using the belief of being first in one's group or overestimation of rank as an indicator does not affect results. Omitting the control for round two performance does not change results either.

[21]Using a dataset containing information on the five highest paid executives in large US corporations for the years 1992-97, Bertrand and Hallock (2001) find that the representation of women reaches a mere 2.5 percent. Also, the gender wage gap is increasing across the wages distribution Arulampalam et al. (2007) and is thus highest for those positions where competition is especially fierce.

or urine samples or to conduct a placebo-controlled trial. The literature on the effects of hormones on economic decision making has strongly focused on testosterone and to a lesser extent on oxytocin, cortisol and oestrogen. Our results suggest the possibility that progesterone could play an equally important role in explaining individual differences in competitiveness and possibly other areas of economic decision making as well. Further research into the exact mechanisms underlying the effects of hormones on competitiveness also seems warranted. This includes the open question of whether it is the preferences of individuals or rather their perceptions of competitive situations which are influenced by hormones.

Chapter 3

Digit Ratios, the Menstrual Cycle and Social Preferences

3.1 Introduction

There is a large literature showing that levels of altruism, trust and reciprocity vary strongly across individuals. While we know that culture plays a role[1], it is unclear whether nature too plays a part in forming these differences. In this paper, we analyse the correlation between behaviour in social preference games and physical proxies for biological factors. In particular, we use the second-to-fourth digit ratio, which is thought to be a marker for prenatal hormone exposure, and menstrual cycle information, which is a proxy for current fluctuations in a range of hormones and other biological processes. We argue that a significant correlation between these markers and economic choices in social situations would suggest that biological factors play a role in shaping social preferences.

The second-to-fourth digit length ratio (2D:4D), the ratio of the length of the index finger to the length of the ring finger, is established in utero and is frequently used as a proxy for prenatal exposure to testosterone and oestrogen which in turn is thought to have a crucial impact on brain development (Hines, 2011). The higher the exposure to testosterone and the lower the exposure to oestrogen, the longer the length of the ring finger relative to the index finger, which leads to a lower 2D:4D. The literature supporting these links is summarised in Section 3.2.1. Millet (2011) summarises the economic literature, which mainly emphasises 2D:4D as a proxy for the relative strength of prenatal testosterone exposure. This literature has focused mainly on risk preferences,

[1] Oosterbeek et al., 2004; Henrich et al., 2001; Gächter et al., 2010; Herrmann et al., 2008.

finding mixed results.[2] But there is a small number of studies looking at behaviour in social preference games, finding a negative correlation of 2D:4D with rejection rates in the ultimatum game in an all-male sample (van den Bergh and Dewitte, 2006) and a negative correlation with giving in the dictator game (Millet and Dewitte, 2006).[3] Sanchez-Pages and Turiegano (2010) find a non-monotonic effect of 2D:4D in a prisoners dilemma whereby subjects with intermediary finger ratios are most likely to cooperate.

The menstrual cycle is characterised by predictable variations in the levels of a range of hormones, as well as other physiological mechanisms such as body temperature. These are described in detail in Section 3.2.2. While this is the first study looking at the correlation between the menstrual cycle and social preferences, a number of previous studies in economics have used the menstrual cycle as a proxy for hormonal fluctuations. Chen et al. (2009) and Pearson and Schipper (2011b) find that bidding in first price auctions fluctuates over the cycle. Buser (2011), reported in Chapter 2 of this thesis, and Wozniak et al. (2010) find cycle effects for competitiveness, a trait which in turn has been found to be related to inequality aversion (Bartling et al., 2009).

A number of studies have also found links between specific hormones and social preferences. A series of placebo controlled studies demonstrates that oxytocin induces higher offers in the trust game (Kosfeld et al., 2005; Baumgartner et al., 2008) and increases generosity in the ultimatum game (Zak et al., 2007). Endogenous oxytocin, stimulated through massage and receiving money in a trust game, increases reciprocity (Morhenn et al., 2008). Burnham (2007) detects a positive correlation between current testosterone levels and rejections in the ultimatum game and Zak et al. (2009), in a placebo controlled study, find testosterone to cause both lower offers and more rejections in men. Conversely, Eisenegger et al. (2010) find that testosterone increases ultimatum offers in women. Randomly treating a sample of post-menopausal women with oestrogen and testosterone, Zethraeus et al. (2009), on the other hand, find no impact on altruism, trust or fairness.[4]

We conduct a laboratory experiment on social preferences and collect information on menstrual

[2] See Dreber and Hoffman (2007); Apicella et al. (2008); Sapienza et al. (2009); Garbarino et al. (2011); Schipper (2011a); Branas-Garza and Rustichini (2011) and Coates et al. (2009). Also see Pearson and Schipper (2011a) on bidding in auctions and Kastlunger et al. (2010) on tax compliance in a laboratory game, both of whom fail to find significant effects of 2D:4D.

[3] However, van den Bergh and Dewitte (2006) find that exposure to sexually stimulating photographs reverses the relationship between 2D:4D and rejections in the ultimatum game and Millet and Dewitte (2009) find that exposure to aggression cues reverses the relationship between 2D:4D and giving in the dictator game.

[4] A number of further studies have investigated the link between biological factors and social preferences using other approaches. Comparing the behaviour of monozygotic and dizygotic twins, a series of studies demonstrates that giving and reciprocity in the trust game (Cesarini et al., 2008), responder behaviour in the ultimatum game (Wallace et al., 2007), and generosity in the dictator game (Cesarini et al., 2009) are partly hereditary. Yet another strand of the literature has found links between specific genes and behaviour in the dictator game (Knafo et al., 2008; Israel et al., 2009). Finally, using the same data used in this study we find significant correlations between handedness and choices in the trust, ultimatum, and dictator games (see Chapter 4).

cycles, contraceptive use, and 2D:4D through a post-experimental questionnaire. This paper contributes to the literature in several ways. We are the first to investigate the correlation between 2D:4D and trust and positive reciprocity. Compared to previous investigations on 2D:4D and other social preferences, our study features a larger range of games and a much larger sample size. This study is also the first to investigate the impact of the menstrual cycle on social preferences, using a range of social preference games covering trust, altruism, and positive as well as negative reciprocity.

We find that digit ratios are strongly and significantly correlated with choices in the social preference games. Subjects with a lower digit ratio give less in all four games. We find at least some evidence that giving in the trust, public good and dictator games, and reciprocity in the trust and ultimatum games vary over the menstrual cycle. The results are most robust for giving and reciprocating in the trust game. These findings are consistent with the hypothesis that social preferences are partially biologically predetermined and are influenced by contemporary biological processes.

Section 3.2 gives details on our biological markers. Section 3.3 describes the data and Section 3.4 explains the experimental design. Section 3.5 describes the results and Section 3.6 discusses possible mechanisms. Section 3.7 concludes.

3.2 Biological markers

3.2.1 2D:4D

The ratio of the length of the index finger to the length of the ring finger (2D:4D), which is established in utero, has been used extensively as a marker for the strength of prenatal hormone exposure (see Manning, 2002, for an introduction). Outside of economics, 2D:4D has been found to be correlated with many traits including reproductive success (Manning et al., 2000), sexual orientation (Robinson and Manning, 2000) and competitiveness in sports (Manning and Taylor, 2001).[5] 2D:4D is thought to be negatively correlated with testosterone exposure – the stronger the exposure, the shorter the index finger relative to the ring finger – and positively with oestrogen exposure in utero. It is reliably higher in women, at least in samples of white individuals. Some studies have also pointed towards a correlation between 2D:4D and current hormone concentrations in adults (Manning et al., 1998) but a meta study (Hönekopp et al., 2007) concludes that there is no significant link. However, lower 2D:4D has also been associated with higher sensitivity to the effects of testosterone (Manning et al., 2003). Higher sensitivity to testosterone amplifies the

[5] See Putz et al. (2004) for a summary of sexually dimorphic traits which have been found to correlate with 2D:4D. In a replication, the authors fail to reproduce most of them.

effects of prenatal testosterone but may also mean that adults with lower 2D:4D are more affected by current hormone levels.

The most direct evidence for the link between 2D:4D and prenatal hormone exposure comes from Lutchmaya et al. (2004) who measure foetal oestrogen and testosterone levels before birth and record digit lengths at age two. They find that the right-hand digit ratio is significantly correlated with prenatal testosterone levels and the ratio of testosterone to oestrogen levels. Earlier studies show that individuals with conditions associated with very high prenatal testosterone levels exhibit significantly smaller 2D:4D (Brown et al., 2002).[6] 2D:4D is determined before birth and thus before economic, social, or cultural factors could shape social preferences.

We collect 2D:4D information through a post-experimental questionnaire by asking subjects the following questions: "On your left (right) hand, which finger is longer: the index finger or the ring finger?" (the possible answers are "index finger", "ring finger", and "equal length"). While this measure is less precise than, for example, hand scans it has a number of advantages. It can be easily included in any questionnaire and is less intrusive on the privacy of subjects. In order for this approach to be valid, there needs to be a strong correlation between the answers to these questions and finger ratios. We validated our measure by taking hand scans from 78 undergraduate students at the University of Amsterdam and asking them the same questions. The difference in average 2D:4D (as measured from the hand scans) between individuals indicating a longer ring finger on both hands, those indicating a longer index finger on both hands and those in between is highly significant (p=0.00; Kruskal Wallis test).[7] The difference in 2D:4D between those indicating a longer ring finger on both hands and those indicating a longer index finger on both hands is equal to approximately 1.2 standard deviations.

A potential issue with using 2D:4D is that digit ratios and the effect of digit ratios on outcomes may vary between individuals of different race (Pearson and Schipper, 2011b). We did not collect the race of our subjects but we collected nationality. In order to reduce noise and eliminate a potential confounding factor, we exclude non-European subjects from the sample we use in our digit ratio regressions. A further potential problem with our measure is that a subject's psychological state may influence the assessment of their own finger length. For example, if subjects are aware of gender differences in digit ratios, those who feel more masculine may be biased towards reporting a longer ring finger. We can directly address this question as in the post-experimental questionnaire, we asked subjects the following question: "On a scale from 1 (very feminine) to 6 (very masculine), how would you describe yourself?". This measure correlates highly with gender (p=0.00; ranksum

[6] See Lutchmaya et al. (2004) for a summary of further indirect evidence for the link between prenatal testosterone and 2D:4D.

[7] The measure is also significant for the right and left hands individually (p=0.00 and p=0.01) and for men and women separately (p=0.05 and p=0.01). The sample of 78 students consisted of 33 men and 45 women.

test) but not with finger ratios or outcomes.[8] Moreover, in our validation sample we find very strong correlations between self-assessed and objectively measured digit ratios for both men and women separately.

3.2.2 The menstrual cycle

The medical literature commonly divides the menstrual cycle into five phases (see e.g. Richardson, 1992) across which the levels of several hormones fluctuate according to a predictable pattern. Assuming a menstrual cycle standardised to 28 days, we can distinguish the menstrual phase (days 1 to 5), the follicular phase (days 6 to 12), the ovulatory phase (days 13 to 15), the luteal phase (days 16 to 23), and the premenstrual phase (days 24 to 28). The levels of oestrogen, progesterone, the luteinising hormone (LH), and the follicle-stimulating hormone (FSH) all fluctuate over the cycle (see Figure 1). Blood levels of oxytocin have also been shown to vary over the menstrual cycle with a peak during the follicular and ovulatory phases (Salonia et al., 2005; Altemus et al., 2001). Testosterone concentration also varies over the cycle, experiencing a mid-cycle peak which coincides with the peak in FSH and LU concentrations (Bloch et al., 1998). Finally, body temperature also varies over the cycle experiencing a sharp increase during ovulation.

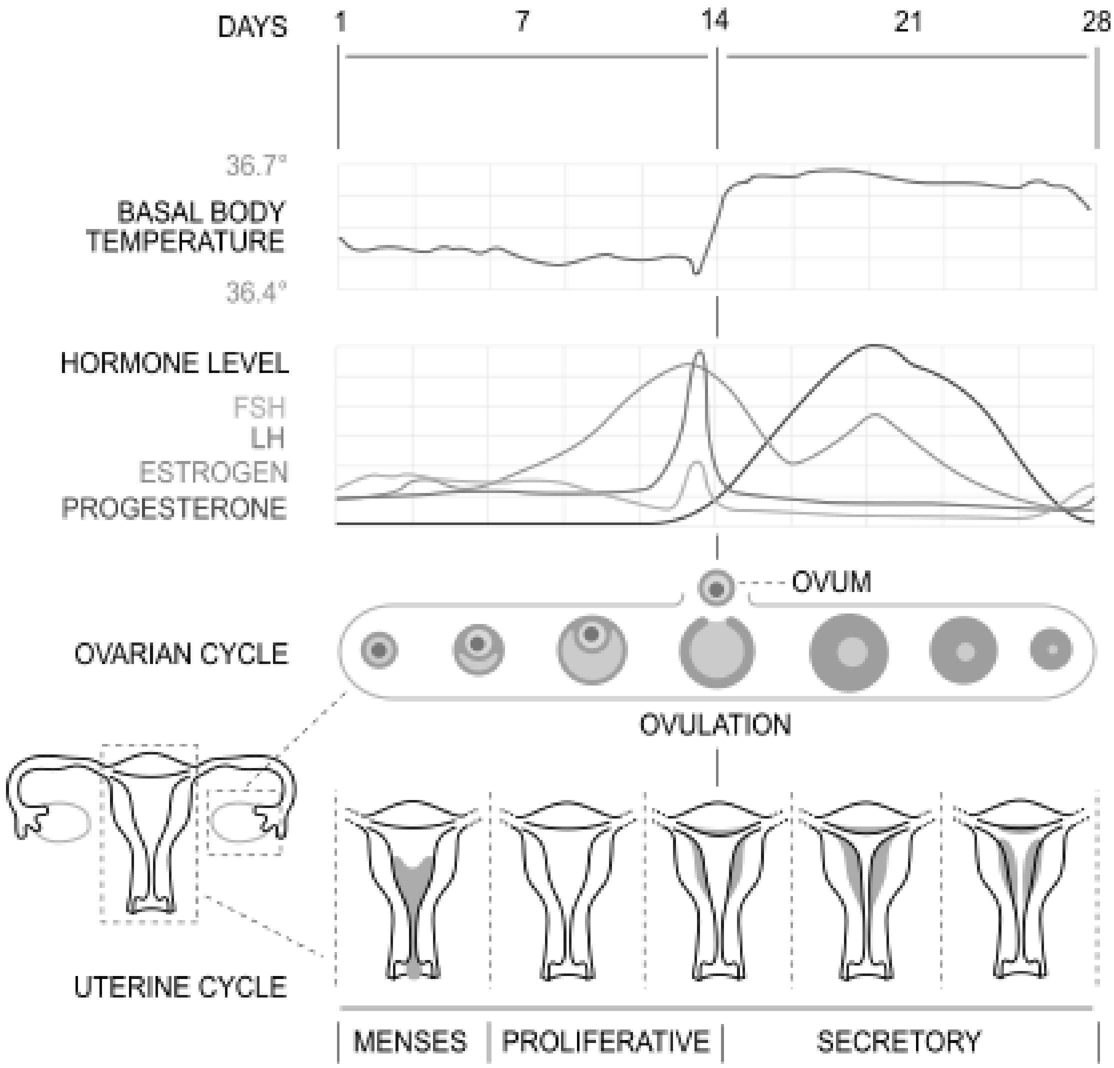
We allocate subjects experiencing a natural menstrual cycle to one of the five menstrual cycle phases based on the cycle information collected through our post-experimental questionnaire. We solicited detailed menstrual cycle information including the beginning of the last menstruation, average cycle length, current menstrual bleeding and regularity of the cycle. Most of the variability in cycle length between individuals stems from differences in the length of the follicular phase. The length of the ovulatory, luteal, and premenstrual phases on the other hand is similar across individuals (Hampson and Young, 2008). We therefore adjust the length of the follicular phase to account for differences in cycle length between subjects.[9] We further reduce misallocations by only allocating subjects to the menstrual phase who indicate to be currently experiencing menstrual bleeding.

Using self-reported menstrual cycle data nevertheless introduces measurement error. Women often misestimate their cycle length (Small et al., 2007) and cycle length tends to vary around the mean

[8] A regression of a ring-finger dummy (indicating the subject reported a longer ring finger on both hands) on self-rated masculinity yields a parameter estimate of 0.036 (s.e. 0.081) for men and -0.036 (0.046) for women. An analogue regression for index fingers yields parameter estimates of -0.007 (0.064) for men and 0.010 (0.047) for women. We also regressed self-rated masculinity on choices in our games using the same set of controls as in Section 3.5. The coefficient is never significant which further confirms that a bias in reported finger ratios due to self-perceived masculinity cannot explain our results.

[9] This method is taken from Pearson and Schipper (2011b).

Figure 3.1: Hormone Levels over the Menstrual Cycle



Source: http://en.wikipedia.org/wiki/File:MenstrualCycle2_en.svg over time (Creinin et al., 2004). Moreover, while the follicular phase varies most with the length of the cycle, there is some variability in the length of other phases too (Stern and McClintock, 1998). As women estimate their cycle length correctly on average (Creinin et al., 2004), this leads to random misallocations.[10] To check whether our results are robust to our adjustment of menstrual cycle phases, we also report *p*-values for two alternative specifications: one where we account for cycle length by adjusting the length of all cycle phases proportionally (universally adjusted phases) and one where we do not adjust phase length at all (unadjusted phases). As a further robustness check, we conduct all our analyses on the subsample of regularly cycling women as well.

Compared to placebo controlled studies on individual hormones, it is difficult to pinpoint precise mechanisms using menstrual cycle information. But contrary to exogenously induced hormone shocks the menstrual cycle occurs naturally and constantly, making the estimated effects especially relevant as predictors of choices in social situations. Moreover, the women in our sample have years of experience with the hormonal fluctuations brought about by the menstrual cycle and therefore a lot of time to adapt to them. An impact of the menstrual cycle on social preferences would therefore constitute strong evidence of an impact of biological factors.

3.2.3 Hormonal contraceptives

In women using hormonal contraceptives such as the pill, vaginal rings or contraceptive patches, which contain varying levels of artificial oestrogen and progestins[11], hormonal fluctuations are different. These contraceptives have in common that they are subject to a 28-day cycle wherein a 21-day intake period, which is characterised by constant daily hormone doses, is followed by a 7-day break during which hormone intake levels drop to zero.[12] Oestrogen excretion by the body is markedly reduced in women taking hormonal contraceptives and progesterone excretion ceases almost completely (Rivera et al., 1999). This leads to a regular pattern whereby levels of artificial oestrogen and progestin are high during the 21-day intake period and low during the 7-day pill break. Oxytocin levels do not vary over the cycle for contraceptive takers (Salonia et al., 2005).

We use our post-experimental questionnaire to ask a series of detailed questions concerning the use of hormonal contraceptives, including the kind of contraceptive and the number of days into the current pill packet or the current pill-break.[13] This allows us to construct a binary variable

[10] Creinin et al. (2004)'s subject pool consists of women reporting a regular menstrual cycle. We also ask subjects whether their cycle is regular and estimate all our regressions also on the sub-sample of regularly cycling subjects.

[11] A progestin is a synthetic hormone that has effects similar to progesterone.

[12] For some contraceptives, the length of the intake phase and break may be slightly different (e.g. 24/4).

[13] Three subjects stated using hormonal contraceptive rings which are placed in the vagina for 21 days followed by a break of 7 days. We asked those subjects how many days ago they applied the ring they are currently using or how many days they are into the current break.

indicating whether a subject currently has high or low hormone levels. The female subjects in our sample take many different oral contraceptives. By comparing subjects on the pill break with those taking the pill, we measure the effect of the average contraceptive. It might seem attractive to use the information on oestrogen dosage and progestin content of the oral contraceptives to disentangle effects. But the exact contraceptive brand prescribed to an individual – and therefore the oestrogen and progestin dosage – is likely correlated with many things including baseline hormone levels and is possibly endogenous with respect to the outcome variables. For the same reason, one cannot derive causal effects from a comparison of pill-takers and naturally cycling women.

3.3 Data

Table 3.1 contains descriptive statistics. As expected, men are significantly more likely to indicate a longer ring finger on both hands (Wilcoxon rank-sum test; p=0.01) while women are significantly more likely to indicate a longer index finger (p=0.00).[14] Table 3.2 shows the sample distribution of digit lengths for European subjects (N=221). 86 subjects have a longer ring finger and 67 have a longer index finger on both hands, with the remaining 68 subjects having mixed ratios.

Table 3.1: Lab sample characteristics

	Sample	Women	Men
N	252	157	95
2x ring finger longer	99	52	47
2x index finger longer	78	60	18
ring finger longer (right)	119	65	54
index finger longer (right)	96	70	26
ring finger longer (left)	121	63	58
index finger longer (left)	97	73	24
Age	22.13	22.02	22.32
Nationality (in %):			
Dutch	71.83	68.79	76.84
Western Europe	3.97	4.46	3.16
Eastern Europe	11.90	13.38	9.47
Asia	8.73	11.46	4.21
America	2.38	1.91	3.16
Other	1.19	0	3.16

[14]This also applies to the right and left hand separately. Gender difference p-values for ring and index finger lengths are 0.018 and 0.007 respectively for the right hand and 0.001 and 0.001 respectively for the left hand.

Table 3.2: Digit ratio distribution for European subjects

	Right 2D:4D			
Left 2D:4D	Longer index	Equal length	Longer ring	Total
Longer index	67	7	11	85
Equal Length	7	15	7	29
Longer ring	9	12	86	107
Total	83	34	104	221

Table 3.3 shows the distribution of female subjects across menstrual cycle phases. Five subjects indicated using implanted contraceptive devices (IUD) which completely suppress menstruation and are therefore excluded from the analysis based on cycle phases. Of the remaining 152 female subjects, 81 do not take hormonal contraceptives and therefore experience a natural cycle. When including subjects whose cycle length is irregular, menstrual phase and follicular phase subjects seem overrepresented. As mentioned above, the follicular phase is the most variable in length and our algorithm therefore adjusts the length of the follicular phase accordingly. Women who experience an irregular cycle – and whose average cycle length is therefore a bad predictor of the current cycle – report cycles that are 3.9 days longer on average (p=0.00) and are therefore more likely to be allocated to the follicular phase. Also, irregularly cycling women may experience intermittent menstrual bleeding and be falsely assigned to the menstrual phase. Indeed we find that 21 out of 30 irregularly cycling women are assigned to phase 1 or 2.[15]

Table 3.3: Subjects by menstrual cycle phase

Menstrual Cycle or Pill Cycle Phase	All subjects	Regular cyclers
Menstrual Phase	17	10
Follicular Phase	31	17
Ovulatory Phase	8	6
Luteal Phase	9	8
Premenstrual Phase	16	10
Total	81	51
Pill Break (7 days)	24	
Pill Intake Phase (21 days)	47	
Total	71	

[15] Menstrual cycle details are sensitive information to ask and a female assistant was therefore present at all sessions and was responsible for all interactions with the subjects concerning the post-experimental questionnaire. In the end, selective non-response turned out not to be a problem as all subjects chose to answer the questions.

3.4 Experimental design

The experiment consists of four social preference games which have been widely used in the literature: a trust game, an ultimatum game, a public good game, and a dictator game. Overall, the experiment lasted for seven rounds and one of the rounds was randomly picked for payment at the end of the experiment. Subjects also received a show-up fee of €10. We ran a total of twelve sessions in December 2009 and January 2010, all of which were conducted in the computer laboratory of the Center for Research in Experimental Economics and Political Decision-Making (CREED) at the University of Amsterdam. The subjects were recruited through CREED's online recruitment system whereby more slots were allocated to female participants to ensure a large enough sample for the analysis of menstrual cycle effects. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). The sessions lasted for approximately two hours and average earnings were around €21.[16]

In the trust game (Berg et al., 1995), two subjects are paired up and each receives an endowment of €10. The first mover (the "Proposer") can then decide how much of his endowment he wishes to send to the second mover (the "Responder"). The amount sent is tripled and the Responder can then decide how much of the money, including his endowment, to send back to the Proposer. Because the Responder has no financial incentive to send back anything, the subgame-perfect Nash equilibrium predicts that the Proposer will not send any money. In the social optimum, on the other hand, the Proposer would send his entire endowment of €10 and the Responder would return less than €30 and more than €10, leaving both parties better off. There is a large literature showing that Proposers send on average around 50 percent of their endowment and that Responders reciprocate by returning on average nearly 50 percent of the received transfer (Levitt and List, 2007).

In the ultimatum game (Güth et al., 1982), the Proposer receives an endowment of €20 while the Responder starts out with nothing. The Proposer decides how much to send to the Responder who can then decide whether to accept or reject the proposal. In case of rejection, both players receive zero, so that the subgame-perfect Nash equilibrium predicts that all positive offers are accepted and Proposers should thus send the lowest possible amount. Again, there is a large literature showing that Proposers send positive amounts, usually in the range between 25 and 50 percent of their endowment, and that Responders are willing to forfeit money by rejecting low offers (Roth, 1995).

The public good game is a generalisation of the prisoner's dilemma game whereby subjects are matched in groups of four and are each endowed with €15. They can then decide how much of the endowment to keep and how much to give to the group. Each Euro given to the group is doubled

[16]The data from this experiment has also been used by Buser (2010), who finds that left-handed men give more in the trust and ultimatum games, are more reciprocal in the trust game, and are less likely to reject offers in the ultimatum game. Left-handed women are less generous in the dictator game.

and split equally amongst the group members such that each Euro given to the group pays 50 Cents to each group member. The social optimum is for all the players to invest everything, but as each player has an incentive to free-ride, the Nash equilibrium predicts zero contributions. There is a large literature reporting substantial positive contributions, usually around 50 percent of the initial endowment in a one-shot setting (Ledyard, 1995).

Finally, we implemented a binary version of the dictator game. In the dictator game, the Proposer again receives an endowment of €20 and has to pick between two options: splitting the pot equally with the Responder (who receives no endowment) or keeping €18 while giving only €2 to the Responder. The Responder has no possibility to reciprocate and the game is consequently a good tool for measuring altruism. The Nash equilibrium of course predicts that the Proposer sends the smallest amount possible, but a large literature finds that when able to decide freely, over 60 percent of subjects send a positive amount (Roth, 1995).

The experiment lasted for seven rounds: two rounds of the ultimatum game, two rounds of the trust game, one round of the public good game, and two rounds of the dictator game (in this order). In all games, only discrete amounts of money could be chosen. For the ultimatum, trust and dictator games each subject played each role exactly once. Subjects were rematched after each round, with the rematching occurring within clusters of eight subjects, and were paired with a different subject in each round.[17] The observations are consequently divided into 34 independent clusters.[18] We take this into account by using clustered standard errors in all regressions.

The fact that all subjects played both roles in each game and played all games in the same order can lead to order effects and experience effects within and between games. The nature of the games makes it impossible to give no feedback at all and we opted for giving full feedback after every round. The effects of experience in previous rounds are in expectation orthogonal to the estimated finger ratio and menstrual cycle effects as subjects are randomly allocated to clusters and rematched across rounds. To further ensure that our results are not due to order effects or past experience, we control for a first mover dummy and past experience in all regressions. This means that in all our regressions, we include controls for the play of opponents in all previous rounds.

[17] With the exception of the public good game which uses groups of four players and in which subjects were matched with at least some players with whom they had previously interacted.

[18] In 5 out of 12 sessions, the number of subjects was not divisible by 8 which leads to 5 clusters which only contain 4 subjects.

3.5 Results

In all following regression tables, the dependent variables are the initial offer and the proportion returned in the trust game in Columns 1 and 2[19], the initial offer in the ultimatum game and a rejection dummy for ultimatum responders in Columns 3 and 4, the contribution to the public good in Column 5, and a binary indicator for choosing the selfish allocation in the dictator game in Column 6. The regressions in Columns 2 and 4, which deal with responder behaviour, control for the amount received from the proposer. Additionally, we control for age and nationality in all regressions.[20] To control for learning and experience effects, standard errors are clustered as described in Section 3.4 and we control for experience in previous rounds in all regressions.

3.5.1 2D:4D

Table 3.4 shows results for regressions of choices in the experimental games on digit ratio indicators for the subsample of European subjects.[21] The first three regressions compare subjects for whom their ring finger is longer than their index finger (low 2D:4D individuals) to the rest of the sample. The rest of the regressions split the sample into three groups, using dummies for a longer ring finger (low 2D:4D) and a longer index finger (high 2D:4D). Subjects with fingers of equal length (intermediate 2D:4D) are therefore the reference group.[22] Table 3.5 summarises the results from these regressions and reports *p*-values for joint significance of digit ratio dummies. All regressions control for gender to avoid that the finger ratio coefficients simply pick up gender effects. Also, qualitative results are the same for both genders separately.[23]

The first specification in Table 3.4 shows that subjects with a longer ring finger on both hands give significantly less than the rest of the sample in the trust, ultimatum, public good and dictator games. Hence, 2D:4D is positively correlated with giving rates in all games. The second specification shows that when using the right hand ratio only (which is the one most commonly used in the literature), subjects with a longer ring finger also return a significantly lower proportion in the trust game but the coefficient on giving in the ultimatum game is now marginally insignificant (p=0.12). All effects go in the same direction when using the left hand ratio only in the third specification

[19]The proportion returned is calculated including the €10 endowment of the responder.

[20]Results without these controls are very similar and are therefore not reported.

[21]Results are very similar when using the whole subject pool.

[22]In the first specification, which uses both hands, the reference group is made up of people who do not have a longer ring or longer index finger on both hands.

[23]The only exceptions are rejection in the ultimatum game, where men with a lower 2D:4D are insignificantly less likely to reject and women with a lower 2D:4D are insignificantly more likely to reject, and contributions to the public good, where women show the same positive correlation of 2D:4D with contributions as the men for the right hand ratio and the double ratio, but a zero correlation for the left hand ratio.

Table 3.4: Digit ratios (2D:4D) and social preferences

	(1) Trust Offer	(2) Proportion Returned	(3) Ultimatum Offer	(4) Ultimatum Rejection	(5) Public Good Contribution	(6) Selfish Dictator
Longer ring	-1.014**	-0.019	-1.008**	0.029	-1.101*	0.076**
(left and right)	(0.395)	(0.013)	(0.414)	(0.044)	(0.593)	(0.035)
Longer ring	-0.869**	-0.038***	-0.719	0.021	-1.504**	0.094**
(right)	(0.397)	(0.012)	(0.450)	(0.047)	(0.556)	(0.039)
Longer ring	-0.968**	-0.028*	-1.106**	0.016	-0.767	0.038
(left)	(0.406)	(0.015)	(0.412)	(0.039)	(0.621)	(0.040)
Longer index	0.463	0.025	0.118	0.025	0.111	0.035
(left and right)	(0.549)	(0.021)	(0.596)	(0.047)	(0.815)	(0.059)
Longer ring	-0.793	-0.007	-0.952**	0.041	-1.048	0.093**
(left and right)	(0.497)	(0.016)	(0.411)	(0.046)	(0.748)	(0.044)
Longer index	-0.171	-0.004	-0.266	0.035	-1.096	0.019
(right)	(0.485)	(0.023)	(0.584)	(0.056)	(0.812)	(0.066)
Longer ring	-0.989*	-0.040**	-0.906*	0.045	-2.273***	0.108*
(right)	(0.501)	(0.018)	(0.476)	(0.063)	(0.777)	(0.060)
Longer index	0.564	-0.008	0.314	0.060	-0.194	0.023
(left)	(0.560)	(0.028)	(0.672)	(0.058)	(1.128)	(0.054)
Longer ring	-0.562	-0.033	-0.879	0.060	-0.907	0.055
(left)	(0.602)	(0.023)	(0.592)	(0.057)	(1.056)	(0.051)
N	221	221	221	221	221	221
Mean	4.172	0.117	7.973	0.136	6.507	0.878
SD	(3.143)	(0.141)	(3.171)	(0.343)	(4.890)	(0.328)
Controls	yes	yes	yes	yes	yes	yes
Offer received	no	yes	no	yes	no	no
Scale	0-10	0-1	0-20	binary	0-15	binary

All coefficients are from OLS regressions. The dependent variables are the initial offer and the proportion returned in the trust game in Columns 1 and 2, the initial offer in the ultimatum game and a rejection dummy for ultimatum responders in Columns 3 and 4, the contribution to the public good in Column 5, and a binary indicator for choosing the selfish allocation in the dictator game in Column 6. Clustered standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls consist of gender, age, nationality and experience in previous games; the regressions in Columns 2 and 4 additionally control for the offer received from the first mover. The regressions are estimated on the subsample of European subjects.

but only trust and ultimatum game giving, as well as trust game reciprocity, remain significant. However, the left-hand ratio is rarely used in the literature and is seen as a less reliable and more noisy predictor of prenatal hormone exposure (see for example Lutchmaya et al., 2004). Rejection rates in the ultimatum game are never significantly correlated with 2D:4D. Our finding of a positive correlation between 2D:4D and giving in the dictator game contradicts the findings of (Millet and Dewitte, 2006).[24]

Table 3.5: Digit ratios (2D:4D) and social preferences: summary of regressions

		(1) Double	(2) Right hand	(3) Left hand
Trust Offer	4D	+(0.015)**	+(0.04)**	+(0.02)**
	2D:4D	+(0.03)**	~(0.10)*	+(0.04)**
Proportion Returned	4D	+(0.16)	+(0.00)***	+(0.07)*
	2D:4D	+(0.20)	+(0.01)**	+(0.16)
Ultimatum Offer	4D	+(0.02)**	+(0.12)	+(0.01)**
	2D:4D	+(0.05)**	~(0.16)	+(0.04)**
Ultimatum Rejection	4D	-(0.52)	-(0.66)	-(0.68)
	2D:4D	~(0.66)	~(0.76)	~(0.54)
Public Good Contribution	4D	+(0.07)*	+(0.01)**	+(0.23)
	2D:4D	~(0.18)	~(0.01)**	~(0.47)
Selfish Dictator	4D	-(0.03)**	-(0.02)**	-(0.34)
	2D:4D	~(0.09)*	-(0.06)*	~(0.53)

*** p<0.01, ** p<0.05, * p<0.1 (p-values in parentheses). + positive effect of 2D:4D (outcome positively correlated with length of index finger and negatively with length of ring finger), - negative effect of 2D:4D (outcome negatively correlated with length of index finger and positively with length of ring finger), ~ non-monotonic effect of 2D:4D (ring and index finger coefficients have the same sign). 4D designates regressions splitting the sample in two groups (ring finger>index finger vs the rest) and 2D:4D designates regressions splitting the sample in three groups (ring finger<index finger, ring finger>index finger, and the rest). All significance levels are from Wald tests for (joint) significance of ring-finger and index-finger coefficients. See Table 3.4 for the regression coefficients.

In the last three specifications, we split the sample into three groups by also adding an index finger dummy. The coefficients therefore now indicate the difference between subjects with a longer ring (or index) finger and subjects for whom ring and index finger have equal length. To determine whether 2D:4D has a significant impact in these regressions, we test for the joint significance of the two finger ratio dummies. The results of these tests are reported in Table 3.5. Most of the results are unaffected – i.e. those with longer ring fingers are less generous than the middle group

[24]The ring finger coefficient for ultimatum rejections is positive, if not significant, which is in accordance with van den Bergh and Dewitte (2006) who find that 2D:4D is negatively correlated with rejection rates.

and those with longer index fingers are more generous. But for the public good game we obtain non-monotonic effects which are significant for the right hand only: subjects with longer ring fingers and subjects with longer index fingers are both less generous than those whose index and ring fingers are of equal length. This is consistent with Sanchez-Pages and Turiegano (2010) who find that individuals with intermediate values of 2D:4D are more likely to cooperate in a prisoner's dilemma game than either low or high 2D:4D subjects.[25]

3.5.2 Menstrual cycle

The graphs in Figure 2 show choices for female subjects in different phases of their menstrual cycle relative to the average choice of male subjects. Table 3.6 reports coefficients for regressions of menstrual cycle phase dummies on choices in all games (men being the reference group). Controls and clustering are as described at the start of Section 3.5. We report coefficients both for regressions using all naturally cycling female subjects and for regressions using only regularly cycling subjects.

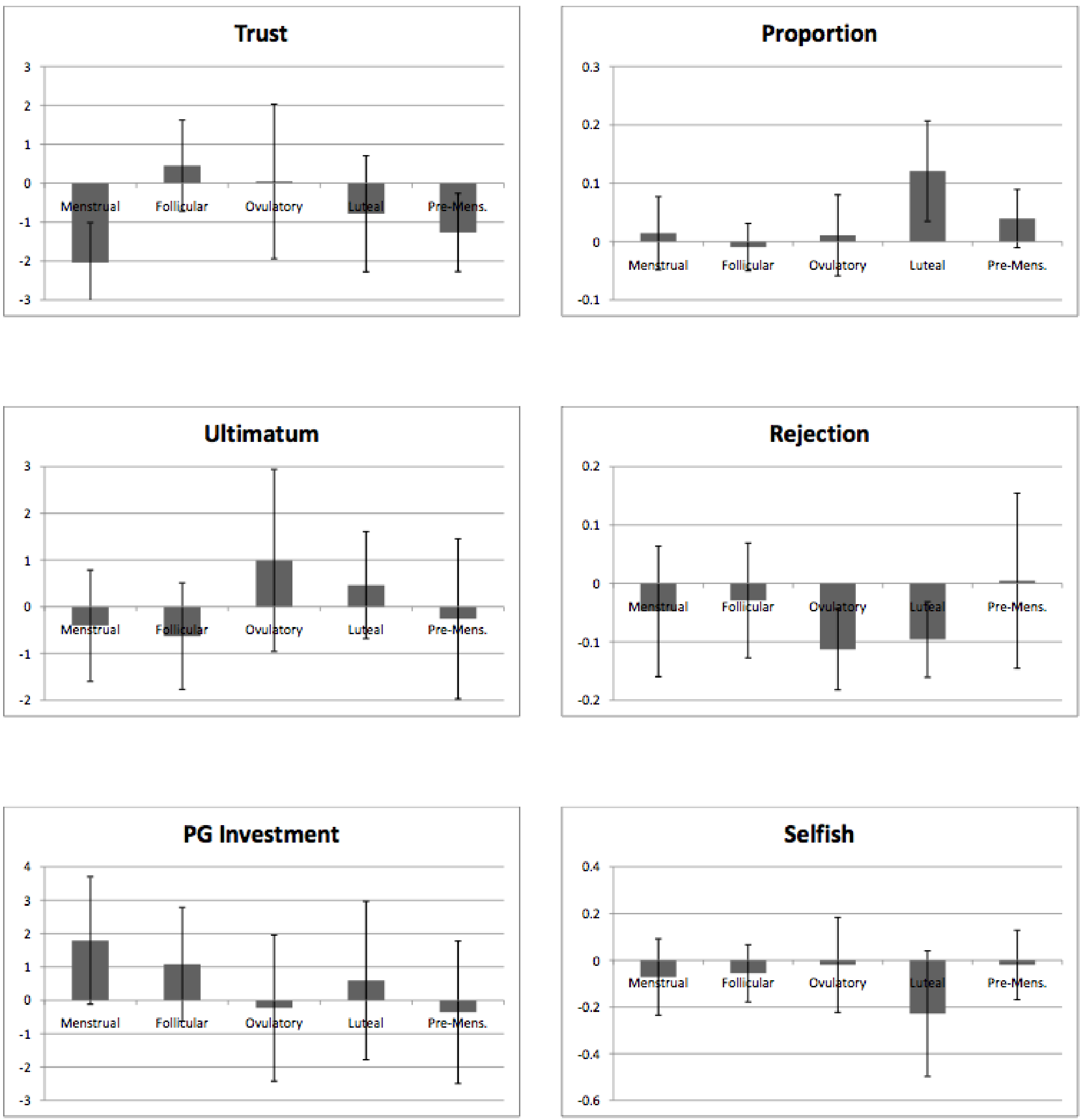
The graphs in Figure 2 show that women give significantly less than men in the trust game during the menstrual and pre-menstrual phases, but not during the middle part of their cycle. Ultimatum game giving also experiences a mid-cycle peak during the ovulatory phase but is never significantly different from the giving rates of men. The likelihood of rejecting ultimatum offers is lower during the ovulatory phase and the luteal phase. Giving in the dictator game and reciprocity in the trust game are highest during the luteal phase. Finally, giving in the public good game is highest during the menstrual phase.

The regression results in Table 3.6 allow us to test which of these fluctuations across the cycle are significant. In each column, we report p-values for a post-estimation test of equality of the five menstrual cycle dummies. The columns with uneven numbers report regressions using the whole sample and the columns with even numbers report regressions using only subjects experiencing a regular cycle. The coefficients on the menstrual cycle dummies represent the difference between the average choices of women in the respective phase and the average choices of men.

The first row of p-values is based on the regressions shown in the table. We can see that when using all naturally cycling women (odd-numbered columns), the variations across the five menstrual cycle phases of trust and public good game giving, as well as the likelihood of rejecting ultimatum game offers, are significant. The variation in trust game reciprocity is marginally significant. We also report p-values using alternative specifications of the menstrual cycle phases as described in

[25] Findings of non-monotonic effects of prenatal testosterone exposure are common; a further example is the impact of prenatal testosterone on homosexuality in men (Robinson and Manning, 2000).

Figure 3.2: The menstrual cycle and social preferences



Coefficients and 90%-confidence intervals are from OLS regressions using the subsample of regularly cycling women; the regressions on trust and ultimatum game reciprocity control for amounts received; effects are shown relative to the average choices of male subjects.

Section 3.2.2. Some of the conclusions change when using universally adjusted or unadjusted phases. Dictator game giving now varies significantly over the cycle but variations in ultimatum game rejections and contributions to the public good are not consistently significant when using these alternative specifications. When we restrict the sample to regularly cycling women only (even-numbered columns), the estimated effects are similar but standard errors increase due to the smaller sample size (51 female subjects). In summary, we find evidence for variations in social preferences across the menstrual cycle but the conclusions are not always robust to the way the menstrual cycle phases are specified. The variations in trust game giving and reciprocity (Columns 1 and 3) appear to be the most robust.

It is difficult to say what these menstrual cycle results predict for contraceptive takers. Many studies which find significant fluctuations of behaviour over the natural menstrual cycle find no effect for contraceptive takers.[26] Moreover, the artificial hormones contained in hormonal contraceptives are not identical to their natural counterparts oestrogen and progesterone. Each pill brand contains a different artificial progestin, some of which, apart from imitating the effects of natural progesterone, also have a list of strong side effects (Mansour, 2006). But Buser (2011), reported in Chapter 2 of this thesis, finds that the effects of hormonal contraceptives on competitiveness are consistent with the difference between the high progesterone luteal phase and the rest of the cycle.[27] This raises the possibility of an impact of hormonal contraceptives on social preferences. The regressions reported in Table 3.7 show that this is not the case. For none of the social preference games are the choices of subjects in the pill-break different from the choices of subjects who were taking contraceptives at the time of the experiment.

3.6 Discussion

We interpret our findings as providing evidence that individual differences in trust, reciprocity and altruism are partially determined by biological factors. Given that our markers proxy for many biological mechanisms, it is difficult to be sure about the exact mechanisms that lay behind our findings. Digit ratios are a proxy for prenatal exposure to both testosterone and oestrogen which in turn are thought to have a multitude of organisational[28] effects. Similarly, a variety of hormones

[26] See Bröder and Hohmann (2003) on risky behaviour, Penton-Voak et al. (1999) on preferences for male faces, Thornhill and Gangestad (1999) on preferences for male body odour, and Miller et al. (2007) on tips earned by lap dancers.

[27] Also, Schipper (2011b) finds that salivary progesterone levels are associated with higher (more competitive) auction bids.

[28] Hormones can have organisational effects as well as immediate effects. Organisational effects are permanent effects of hormones on the structure and functioning of the body and brain. The critical periods during which organisational effects mostly occur are foetal development and puberty.

Table 3.6: Menstrual cycle and social preferences

	(1) Trust	(2) Trust	(3) Proportion	(4) Proportion	(5) Ultimatum	(6) Ultimatum	(7) Rejection	(8) Rejection	(9) PG	(10) PG	(11) Selfish	(12) Selfish
Menstrual	-1.578***	-1.600**	0.034	0.061	-0.070	1.405	-0.058	0.017	2.665*	3.995***	-0.102	-0.094
	(0.460)	(0.589)	(0.037)	(0.042)	(0.847)	(0.839)	(0.064)	(0.086)	(1.358)	(1.427)	(0.101)	(0.136)
Follicular	0.353	-0.546	0.008	-0.031	-0.389	0.431	-0.026	-0.082	1.118	0.779	-0.044	0.050
	(0.641)	(0.683)	(0.026)	(0.030)	(0.790)	(0.925)	(0.062)	(0.081)	(0.773)	(0.878)	(0.076)	(0.086)
Ovulatory	0.494	0.309	0.009	0.009	0.692	2.075*	-0.131**	-0.157**	-1.444	-1.226	0.094	0.138
	(1.085)	(1.464)	(0.041)	(0.048)	(1.282)	(1.175)	(0.061)	(0.077)	(0.957)	(1.172)	(0.137)	(0.202)
Luteal	-0.879	-0.956	0.134**	0.121*	0.727	1.226	-0.083	-0.070	0.875	0.673	-0.210	-0.258
	(0.805)	(0.938)	(0.053)	(0.065)	(0.920)	(0.781)	(0.060)	(0.084)	(1.194)	(1.114)	(0.153)	(0.165)
Pre-mens.	-0.578	-1.120*	0.077**	0.056	-0.096	1.699	0.045	-0.023	1.016	-0.113	-0.083	-0.097
	(0.514)	(0.550)	(0.034)	(0.040)	(1.039)	(1.110)	(0.099)	(0.106)	(1.215)	(1.148)	(0.101)	(0.157)
N	176	146	176	146	176	146	176	146	176	146	176	146
Joint significance:												
P-val.	0.08	0.46	0.10	0.16	0.72	0.80	0.08	0.52	0.03	0.06	0.36	0.16
P-val. universally adj.	0.01	0.26	0.02	0.10	0.77	0.31	0.35	0.11	0.38	0.09	0.01	0.00
P-val. unadjusted	0.06	0.00	0.12	0.00	0.93	0.00	0.29	0.47	0.37	0.00	0.00	0.09
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Offer	no	no	yes	yes	no	no	yes	yes	no	no	no	no
Scale	0-10	0-10	0-1	0-1	0-20	0-20	binary	binary	0-15	0-15	binary	binary

All coefficients are from OLS regressions. The dependent variables are the initial offer and the proportion returned in the trust game in Columns 1 and 2, the initial offer in the ultimatum game and a rejection dummy for ultimatum responders in Columns 3 and 4, the contribution to the public good in Column 5, and a binary indicator for choosing the selfish allocation in the dictator game in Column 6. Clustered standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Controls consist of age, nationality and experience in previous games; the regressions in Columns 3, 4, 7 and 8 additionally control for the offer received from the first mover. The sample consists of women experiencing a natural menstrual cycle and all men (the regressions in even-numbered columns restrict the sample to regular cyclers and men). Men are the reference category.

Table 3.7: Hormonal contraceptives and social preferences

	(1) Trust Offer	(2) Proportion Returned	(3) Ultimatum Offer	(4) Ultimatum Rejection	(5) Public Good Contribution	(6) Selfish Dictator
Pill Break	-0.111	-0.025	0.194	-0.035	0.530	-0.008
	(0.949)	(0.023)	(1.092)	(0.100)	(1.389)	(0.070)
N	71	71	71	71	71	71
Mean	3.647	0.117	8.845	0.197	6.197	0.915
SD	2.747	0.134	3.500	0.401	4.496	0.280
Controls	yes	yes	yes	yes	yes	yes
Offer received	no	yes	no	yes	no	no
Scale	0-10	0-1	0-20	binary	0-15	binary

All coefficients are from OLS regressions. The dependent variables are the initial offer and the proportion returned in the trust game in Columns 1 and 2, the initial offer in the ultimatum game and a rejection dummy for ultimatum responders in Columns 3 and 4, the contribution to the public good in Column 5, and a binary indicator for choosing the selfish allocation in the dictator game in Column 6. Clustered standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls consist of age, nationality and experience in previous games; the regressions in Columns 2 and 4 additionally control for the offer received from the first mover. The sample is restricted to female contraceptive takers.

and other biological factors vary over the menstrual cycle. In this section, we nevertheless briefly discuss possible mechanisms, an endeavour which necessarily remains speculative.

The organisational effects of testosterone and oestrogen exposure occur between weeks seven and twelve of pregnancy and the relative strength of exposure is thought to have a strong impact on preferences and behaviour later in life (see Hines, 2009 for an introduction). It is thought that hormones exert these permanent effects by altering neural structures although the exact mechanisms are still largely unidentified (Hines, 2011). Amongst many identified behavioural effects, prenatal testosterone exposure is thought to have a negative impact on empathy (Baron-Cohen et al., 2004). These differences are present already at birth (Lutchmaya and Baron-Cohen, 2002), and at four years of age children with higher exposure have significantly lower social skills and more restricted interests (Knickmeyer et al., 2005). Empathy has been shown to be positively correlated with giving in social preference games (Barraza and Zak, 2009). A link between prenatal testosterone exposure and empathy is therefore a potential mechanism behind our finding of a positive correlation of 2D:4D with social preferences. However, current testosterone levels in adults have also been associated with social preferences and adults with a lower 2D:4D may be more sensitive to the effects of testosterone (Manning et al., 2003). Although 2D:4D is not significantly correlated with testosterone levels in adults (Hönekopp et al., 2007), the increased sensitivity to testosterone in individuals with low 2D:4D is another potential explanation for our findings.

Trust game giving is highest in mid-cycle during the follicular and ovulatory phases. Blood levels of oxytocin, which has been frequently linked to choices in the trust game, also experience a peak during these phases. However, oxytocin in the blood stream does not permeate the blood-brain-barrier (Ermisch et al., 1985). The correlation is therefore more likely due to the same brain mechanisms that are responsible for the oxytocin release influencing trust game giving rather than to direct action of the oxytocin itself. Testosterone secretion also experiences a mid-cycle peak. But the only study we are aware of that tests the impact of testosterone on trust game giving (Zethraeus et al., 2009) does not find an effect and testosterone has been shown to decrease self-rated trust towards strangers (Bos et al., 2010). Eisenegger et al. (2010) find that testosterone increases ultimatum game giving in women which is consistent with testosterone explaining our finding of a (not statistically significant) mid-cycle peak. LH and FSH also peak in mid-cycle (also coinciding with an increase in body temperature) and it is therefore difficult to nail down the exact mechanism.

Buser (2011), reported in Chapter 2 of this thesis, finds that competitiveness is significantly lower during the luteal phase and is negatively correlated with expected progesterone levels. Bartling et al. (2009) find a negative relationship between competitiveness and preferences for egalitarian choices (choices that reduce favourable or unfavourable payoff inequality). Proposers in the dictator and ultimatum games and responders in the trust game make their choices in situations which are characterised by favourable inequality. The proportion returned in the trust game and giving in the ultimatum game are higher and the likelihood of being selfish in the dictator game is lower during the luteal phase although the difference is not significant in the case of the dictator and ultimatum games.[29] Conversely, we do not find this effect for giving in the trust and public good games where endowments are equal. Progesterone peaks during the luteal phase and a positive effect of progesterone on inequality aversion is therefore a possible explanation for our findings. On the other hand, Buser (2011) finds hormonal contraceptives – which contain artificial progestins – to also have a negative impact on competitiveness whereas we do not find an effect of contraceptives in any game.

[29] OLS regressions of outcomes on luteal phase dummy using sample of naturally cycling women and controlling for age and nationality. Proportion returned in the trust game: p=0.01; ultimatum game offer: p=0.20; likelihood of being selfish in dictator game: p=0.11.

3.7 Conclusions

Our results demonstrate that biological factors, both prenatal and current, play an important role in shaping social preferences. Subjects with a lower 2D:4D are less trusting, less reciprocal and less altruistic. As 2D:4D is established before birth, this is evidence that social preferences are partially biologically predetermined. 2D:4D is, amongst other things, a proxy for prenatal exposure to testosterone which in turn has important organisational effects on the brain. Our results therefore suggest the possibility of a neural basis for social preferences.

For women who experience a natural menstrual cycle, trust, reciprocity and altruism vary across cycle phases. Trust is highest before and during ovulation when several hormones are peaking. Altruism in games characterised by favourable income inequality is highest during the luteal phase which is characterised by a peak in progesterone levels. This points to an increase in inequality aversion at that point of the cycle, possibly mediated by progesterone. However, we encounter zero findings for contraceptive takers. Also, not all of the estimated fluctuations over the cycle are robust to alternative specifications of the menstrual cycle phases and further research with larger sample sizes could confirm the robustness of our results. Nevertheless, the observed variations in social preferences over the cycle provide further evidence for a biological basis for social preferences, specifically pointing towards an impact of hormones.

The biological markers we use are imprecise devices but they have unique advantages. Our 2D:4D indicators are simple enough to be included in questionnaires which is impossible for hand-scans and hormone assays, let alone for direct prenatal hormone measurements during pregnancy. Furthermore – contrary to exogenously induced hormone shocks – the menstrual cycle occurs naturally and constantly, making the estimated effects especially relevant for our understanding of choices in social situations. Our results fit well with the growing literature showing that the strong individual differences observed in a large number of social preference experiments are at least partially biologically determined.

Chapter 4

Handedness Predicts Social Preferences: Evidence Connecting the Lab to the Field

4.1 Introduction

There is a large literature showing that levels of altruism, trust and reciprocity vary strongly across individuals. While we know that culture plays a role[1], it is unclear whether nature too plays a part in forming these differences. In this paper, we analyse the correlation between social preferences and handedness, a trait which is related to important neurological differences. We argue that a significant correlation between handedness and economic choices in social situations would indicate that nature plays an important role in shaping social preferences.

Handedness data can be collected through surveys which allows us to examine the biological origins of altruism, trust and reciprocity in the field as well as the lab. In a first step, we estimate the correlation between handedness and choices in a range of social preference games in the lab. In a second step, we use survey data on altruistic behaviour in the field and attitude questions concerning trust and reciprocity. We can therefore test the external validity of our lab findings by investigating whether they apply to behaviour outside the lab using diverse and representative samples.

Left-handedness is associated with important differences in brain development. For example, while language ability is controlled by the left side (hemisphere) of the brain in 97 percent of right-handers, it is bilateral or controlled by the right side in more than 30 percent of left-handers (Knecht et al., 2000). The brains of left-handers are found to exhibit lower rates of brain hemisphere specialisation in general, meaning that left-handers more commonly use both sides of the brain

[1] Oosterbeek et al., 2004; Henrich et al., 2001; Gächter et al., 2010; Herrmann et al., 2008.

for a given task (Coren, 1993). The main connection between the two hemispheres of the brain (a thick band called corpus callosum, which contains millions of nerves and acts as a data-wire that allows the two hemispheres to "speak" to each other) is thicker in left-handers (Witelson, 1985).[2] It is important to note that this correlation between handedness and the size of the corpus callosum is much stronger in men than in women (Witelson and Goldsmith, 1991; Grimshaw et al., 1995).

Handedness has also been associated with the strength of natural prenatal exposure to testosterone, which is itself thought to have a strong impact on brain development. Prenatal exposure to testosterone has a negative impact on the size of the corpus callosum (Witelson, 1991) and may therefore shift handedness towards the right (Witelson and Nowakowski, 1991). This effect is probably exclusive to men for whom the link between handedness and the size of the corpus callosum is much stronger and who are exposed to much higher levels of prenatal testosterone exposure.[3] The opposite relation (prenatal testosterone being a cause of left-handedness) has also been hypothesised (Geschwind and Galaburda, 1987). The empirical evidence is inconclusive, with some evidence found in both directions.[4]

The psychological literature has found left-handers to be different from right-handers along many dimensions. Often, the effects are stronger in men and it is common to even find opposite effects of handedness for women. Coren (1995), for instance, finds creativity (specifically "divergent thinking") to be positively correlated with left-handedness in men only. Goldberg et al. (1994) find that left-handedness is associated with novelty-seeking, again particularly in males. Looking at several measures of school performance and leadership skills, Faurie et al. (2006) find a positive association with left-handedness for boys but a negative one for girls and Sanders et al. (1982) find the same pattern for spatial ability.

In economics, left-handedness has been associated with higher wages for men but lower wages for women (Denny and O'Sullivan, 2007), higher wages amongst college-educated men (Ruebeck et al., 2007), and worse average outcomes in a range of early childhood development indicators (Johnston et al., 2007). In an experimental study, Hoffman and Gneezy (2010) find that left-handers of both sexes are more competitive. Given that brain differences associated with handedness differ between the sexes and given the findings of the psychological literature, it comes as no surprise that economic studies also find gender differences in the effects of handedness.

A number of previous papers have investigated the link between nature and social preferences in the lab using different methods. One strand of this literature has focused on hormones. A series of

[2] Indeed, it has been hypothesised that handedness arises from differences in the size of the corpus callosum (Witelson, 1985).

[3] Lust et al. (2010) show that higher prenatal testosterone exposure reduces information transfer via the corpus callosum in male subjects only. See Grimshaw et al. (1995) for a summary of the theories (and supporting empirical evidence) linking prenatal testosterone exposure to handedness.

[4] See for example Manning and Peters (2009) or Grimshaw et al. (1995).

placebo controlled studies demonstrates that oxytocin induces higher offers in the trust game (Kosfeld et al., 2005; Baumgartner et al., 2008) and increases generosity in the ultimatum game (Zak et al., 2007). Endogenous oxytocin, stimulated through massage and receiving money in a trust game, increases reciprocity (Morhenn et al., 2008). Burnham (2007) detects a positive correlation between current testosterone levels and rejections in the ultimatum game and Zak et al. (2009), in a placebo controlled study, find testosterone to cause both lower offers and more rejections in men. Conversely, Eisenegger et al. (2010) find that testosterone increases ultimatum offers in women. Randomly treating a sample of post-menopausal women with oestrogen and testosterone, Zethraeus et al. (2009), on the other hand, find no impact on altruism, trust or fairness.

Comparing the behaviour of monozygotic and dizygotic twins, a series of studies demonstrates that giving and reciprocity in the trust game (Cesarini et al., 2008), responder behaviour in the ultimatum game (Wallace et al., 2007), and generosity in the dictator game (Cesarini et al., 2009) are partly hereditary. Yet another strand of the literature has found links between specific genes and behaviour in the dictator game (Knafo et al., 2008; Israel et al., 2009).

In a companion paper using the same data (Buser, 2012, reported in Chapter 3), we correlate the experimental choices in the social preference games with physical markers for the strength of prenatal testosterone exposure, finding that individuals of both sexes with markers indicating stronger prenatal exposure show lower levels of altruism, trust, and positive reciprocity.[5] Moreover, we find some evidence that for female subjects offers and responses in the trust, ultimatum and dictator games vary over the menstrual cycle.

We let a sample of 252 subjects participate in a series of social preference games in the lab, including the standard dictator, ultimatum, public good and trust games. We use a post-experimental questionnaire to collect information on handedness and other demographic variables. We also gather handedness information from the respondents of the LISS internet panel[6] and link it to survey data measuring altruistic behaviour and trust outside of the lab. We also use US data from the National Longitudinal Survey of Youth (NLSY)[7], which contain both handedness information and some indicators of altruism, to further confirm our results.

The rest of the paper is organised as follows. Section 4.2 gives details on the experimental design and the lab data and Section 4.3 describes the experimental results. Section 4.4 describes the survey data and Section 4.5 contains the results obtained using the survey data. Section 4.6 discusses potential biological mechanisms behind our results and Section 4.7 concludes.

[5] Specifically, we use the ratio of the length of the index finger to the length of the ring finger (2D:4D) which is established in utero and is negatively correlated with the strength of prenatal exposure to testosterone in both sexes (Hines, 2011). Further studies correlating 2D:4D with social preferences include van den Bergh and Dewitte (2006), Millet and Dewitte (2006) and Millet and Dewitte (2009).

[6] The LISS panel is a representative Dutch internet panel (www.lissdata.nl).

[7] http://www.nlsinfo.org/

4.2 Experimental design and data

The experiment consists of four social preference games which have been widely used in the literature: a trust game, an ultimatum game, a public good game, and a dictator game. Overall, the experiment lasted for seven rounds and one of the rounds was randomly picked for payment at the end of the experiment. Subjects also received a show-up fee of €10. We ran a total of twelve sessions in December 2009 and January 2010, all of which were conducted in the computer laboratory of the Center for Research in Experimental Economics and Political Decision-Making (CREED) at the University of Amsterdam. The subjects were recruited through CREED's online recruitment system. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). The sessions lasted for approximately two hours and average earnings were around €21.

In the trust game (Berg et al., 1995), two subjects are paired up and each receives an endowment of €10. The first mover (the "Proposer") can then decide how much of his endowment he wishes to send to the second mover (the "Responder"). The amount sent is tripled and the Responder can then decide how much of the money, including his endowment, to send back to the Proposer. Because the Responder has no financial incentive to send back anything, the subgame-perfect Nash equilibrium predicts that the Proposer will not send any money. In the social optimum, on the other hand, the Proposer would send his entire endowment of €10 and the Responder would return less than €30 and more than €10, leaving both parties better off. There is a large literature showing that Proposers send on average around 50 percent of their endowment and that Responders reciprocate by returning on average nearly 50 percent of the received transfer (Levitt and List, 2007).

In the ultimatum game (Güth et al., 1982), the Proposer receives an endowment of €20 while the Responder starts out with nothing. The Proposer decides how much to send to the Responder who can then decide whether to accept or reject the proposal. In case of rejection, both players receive zero, so that the subgame-perfect Nash equilibrium predicts that all positive offers are accepted and Proposers should thus send the lowest possible amount. Again, there is a large literature showing that Proposers send positive amounts, usually in the range between 25 and 50 percent of their endowment, and that Responders are willing to forfeit money by rejecting low offers (Roth, 1995).

The public good game is a generalisation of the prisoner's dilemma game whereby subjects are matched in groups of four and are each endowed with €15. They can then decide how much of the endowment to keep and how much to give to the group. Each Euro given to the group is doubled and split equally amongst the group members such that each Euro given to the group pays 50 Cents to each group member. The social optimum is for all the players to invest everything, but as each player has an incentive to free-ride, the Nash equilibrium predicts zero contributions. There is a

Table 4.1: Lab sample characteristics

	Sample	Women	Men
N	252	157	95
Left-handed	20	11	9
Age	22.1	22.0	22.3
Dutch	72%	69%	77%

large literature reporting substantial positive contributions, usually around 50 percent of the initial endowment in a one-shot setting (Ledyard, 1995).

Finally, we implemented a binary version of the dictator game. In the dictator game, the Proposer again receives an endowment of €20 and has to pick between two options: splitting the pot equally with the Responder (who receives no endowment) or keeping €18 while giving only €2 to the Responder. The Responder has no possibility to reciprocate and the game is consequently a good tool for measuring altruism. The Nash equilibrium of course predicts that the Proposer sends the smallest amount possible, but a large literature finds that when able to decide freely, over 60 percent of subjects send a positive amount (Roth, 1995).

The experiment lasted for seven rounds: two rounds of the ultimatum game, two rounds of the trust game, one round of the public good game, and two rounds of the dictator game (in this order). In all games, only discrete amounts of money could be chosen. For the ultimatum, trust and dictator games each subject played each role exactly once. Subjects were rematched after each round, with the rematching occurring within clusters of eight subjects, and were paired with a different subject in each round.[8] The observations are consequently divided into 34 independent clusters.[9] We take this into account by using clustered standard errors in all regressions.

The fact that all subjects played both roles in each game and played all games in the same order can lead to order effects and experience effects within and between games. The nature of the games makes it impossible to give no feedback at all and we opted for giving full feedback after every round. The effects of experience in previous rounds are in expectation orthogonal to the estimated effects of handedness as subjects are randomly allocated to clusters and rematched across rounds. To further ensure that our results are not due to order effects or past experience, we control for a first mover dummy and past experience in all regressions. This means that in all our regressions, we include controls for the play of opponents in all previous rounds.

[8] With the exception of the public good game which uses groups of four players and in which subjects were matched with at least some players with whom they had previously interacted.

[9] In 5 out of 12 sessions, the number of subjects was not divisible by 8 which leads to 5 clusters which only contain 4 subjects.

The laboratory sample consists of 252 undergraduate students, 157 of whom are female. After participating in the experiment, the subjects also filled out a questionnaire asking whether they are left- or right-handed. We also collected additional demographic information including gender, age and nationality. The responses are summarised in Table 4.1. In our sample 9.5 percent of the men and 7.0 percent of the women are left-handed which is consistent with the literature on handedness which reports a higher prevalence of left-handedness amongst men (McManus, 2002).

4.3 Laboratory results

Table 4.2 contains average choices in the social preference games by handedness and gender. The observed amounts are largely within the range observed in the previous literature, with mean offers in the trust, ultimatum and public good games all being around 40 percent of the endowment. The low proportion of equal allocations in our binary dictator game is also in line with the literature as offers of 50 percent of the endowment are a relatively rare occurrence. Looking at the whole sample, the strongest differences between left and right-handers occur in the trust game, where left-handers give 57 percent more, and in the ultimatum game, where left-handers virtually never reject an offer compared to 13.4 percent of right-handers.

Splitting the sample by gender, it becomes apparent that the effects of handedness differ strongly between men and women. Compared to right-handed men, left-handed men give 86 percent more in the trust game and 26 percent more in the ultimatum game. They also return a higher proportion in the trust game and are much less likely to reject in the ultimatum game. These differences are much weaker in women apart from the difference in the likelihood of rejecting ultimatum game offers. On the other hand, left-handed women are 14 percentage points more likely to be selfish in the dictator game.

We further explore these differences using OLS regressions, the results of which are shown in Table 4.3. The dependent variables are the initial offer and the proportion returned in the trust game in Columns 1 and 2, the initial offer in the ultimatum game and a rejection dummy for ultimatum responders in Columns 3 and 4, the contribution to the public good in Column 5, and a binary indicator for choosing the selfish allocation in the dictator game in Column 6. Using regression analysis allows us to properly control for past experience and take into account clustering, as described in Section 4.2. The regressions in Columns 2 and 4, which deal with responder behaviour, additionally control for the amount received from the proposer. Also, in the regressions using the whole sample we control for gender.

The whole-sample regressions confirm that left-handers give significantly more in the trust game and are less likely to reject an offer in the ultimatum game. Striking differences appear when we

Table 4.2: Descriptive results

	Whole sample:		Men:		Women:	
	left-h.	right-h.	left-h.	right-h.	left-h.	right-h.
Trust offer	6.050	3.845	7.667	4.116	4.727	3.685
	(3.268)	(3.002)	(2.872)	(3.320)	(3.069)	(2.798)
Proportion returned	0.908	0.752	0.995	0.678	0.841	0.795
	(0.641)	(0.716)	(0.542)	(0.697)	(0.735)	(0.726)
Ultimatum offer	8.650	7.935	9.556	7.570	7.909	8.151
	(3.422)	(3.160)	(3.575)	(2.601)	(3.270)	(3.437)
Ultimatum rejection	0.050	0.134	0.000	0.105	0.091	0.151
	(0.224)	(0.341)	(0.000)	(0.308)	(0.302)	(0.359)
PG contribution	6.400	6.228	6.444	5.791	6.364	6.486
	(5.413)	(4.784)	(5.747)	(5.404)	(5.409)	(4.377)
Selfish dictator	0.950	0.875	0.889	0.895	1.000	0.863
	(0.224)	(0.331)	(0.333)	(0.308)	(0.000)	(0.345)

Table 4.3: Handedness and social preferences

	(1) Trust Offer	(2) Proportion Returned	(3) Ultimatum Offer	(4) Ultimatum Rejection	(5) Public Good Contribution	(7) Selfish Dictator
Whole sample:						
Left-handed	2.348***	0.112	0.796	-0.113**	0.139	0.044
	(0.724)	(0.172)	(0.824)	(0.048)	(1.118)	(0.056)
N	252	214	252	252	252	252
Men:						
Left-handed	3.864***	0.359	2.249*	-0.146***	-0.018	0.013
	(1.295)	(0.258)	(1.103)	(0.050)	(1.731)	(0.122)
N	95	80	95	95	95	95
Women:						
Left-handed	1.049	-0.044	-0.187	-0.071	0.180	0.071**
	(0.909)	(0.214)	(1.029)	(0.070)	(1.447)	(0.034)
N	157	134	157	157	157	157
Controls	yes	yes	yes	yes	yes	yes
Offer received	no	yes	no	yes	no	no
Scale	0-10	0-1	0-20	binary	0-15	binary

All coefficients are from OLS regressions. The dependent variables are the initial offer and the proportion returned in the trust game in Columns 1 and 2, the initial offer in the ultimatum game and a rejection dummy for ultimatum responders in Columns 3 and 4, the contribution to the public good in Column 5, and a binary indicator for choosing the selfish allocation in the dictator game in Column 6. Clustered standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls consist of gender, nationality and experience in previous games; the regressions in Columns 2 and 4 additionally control for the offer received from the first mover. The regressions in Column 2 only use subjects who received a positive amount from their proposer.

split the sample by gender. Amongst male subjects, the left-handers make offers in the trust and ultimatum games which are both statistically and economically significantly higher. They are also significantly less likely to reject offers in the ultimatum game. While they return a substantially larger proportion of the money received in the trust game, this difference is not statistically significant. Female left-handers show none of these effects but are significantly less likely than other female subjects to choose the generous allocation in the dictator game. At first glance, these results indicate a positive impact of left-handedness on giving rates for men and a negative impact for women.

We first take a closer look at giving behaviour. Relative to their right-handed counterparts, the male left-handers make choices which leave their partners better off in both the trust and ultimatum games. But in the public good and dictator games, where recipients cannot reciprocate, their choices are indistinguishable from those of right-handers. Higher altruism in left-handers is therefore not the explanation. A possible explanation is that left-handed men are indeed more trusting, expecting others to exercise positive reciprocity in the trust game, and expect others to exercise negative reciprocity in the ultimatum game. That is, left-handed men have a stronger belief in the reciprocity and fairness of others.[10]

In Table 4.4, we take a closer look at the link between handedness and reciprocity in the trust game. Here we regress the amount returned in the trust game on a handedness dummy, the amount received from the proposer, and the interaction of these two variables. It becomes apparent that left-handed men do not return more money per se; actually, the coefficient on the handedness dummy is negative and insignificant. Rather, the amount they return increases much more steeply with the amount received from their proposer. While right-handed men return around 85 cents for each Euro sent by the proposer (or around 28 cents for each Euro they receive after the amount is tripled), left-handed men return almost twice as much per Euro. Given that we do not use the strategy method, the amount of rejections in the data is too small in order to conduct the same analysis for the ultimatum game. But in the light of these results, it seems plausible that for left-handed men the willingness to reject reduces faster with an increase in the transfer received and that this leads to the significantly lower likelihood of rejection. These results again indicate that the differences in behaviour between left-handed and right-handed men are not driven by differences in altruism (in that case the left-handers should return more whatever the amount received) but by a stronger willingness to reciprocate.

[10] Similarly, left-handed men might simply be better at inferring the reactions of others so that their higher giving rates translate into higher profits. This is not the case as ultimatum proposer profits are virtually the same for left-handed and right-handed males and trust proposer profits are (insignificantly) lower for left-handers. On the other hand, the ultimatum offers of left-handed men are 12.8 percent less likely to get rejected than those of right-handed men (OLS regression without additional controls; p=0.001).

Table 4.4: Handedness and reciprocity in the trust game

	(1) Amount returned	(2) Amount returned	(3) Amount returned
	Full Sample:	**Men:**	**Women:**
Left-handed	-0.535	-0.918	-1.017
	(0.928)	(1.038)	(1.638)
Offer received	1.047***	0.931***	1.135***
	(0.114)	(0.253)	(0.124)
Left-handed *x* Offer received	0.266	0.772**	0.297
	(0.260)	(0.350)	(0.352)
N	252	95	157
Mean	3.339	2.758	3.771
SD	4.792	4.346	5.018
Controls	yes	yes	yes

All coefficients are from OLS regressions. Clustered standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls consist of nationality, gender and experience in previous games.

The female sample shows a strikingly different pattern. There are no significant differences in giving between right-handed and left-handed women in either the trust or the ultimatum game. The same is true for the proportion returned in the trust game and the likelihood of rejection in the ultimatum game. The regression results reported in Column 3 of Table 4.4 show that there are no differences in conditional reciprocity in the trust game either. On the other hand, left-handed women are significantly more likely to be selfish in the dictator game where reciprocity plays no role. Given that the neurological differences associated with handedness differ between men and women, it comes as no surprise that the effects of handedness on social preferences vary between the genders.

This section presented experimental results which imply that the differences in neural structures associated with handedness play a role in shaping individual differences in social preferences. The observed effects of handedness are, for instance, stronger than most of the cultural differences or gender effects observed in the social preference literature.[11] In the following sections, we will investigate whether these links between handedness and social preferences are present in altruistic and reciprocal behaviour outside of the lab as well.

[11] For cultural differences in social preferences refer to the literature mentioned above. For a review of gender differences in social preferences see Croson and Gneezy (2009).

4.4 Survey data

An often voiced concern about laboratory experiments measuring social preferences is that their findings might not be externally valid. Levitt and List (2007) identify a range of potential problems with generalising the results of lab experiments on social preferences including the effect of scrutiny by the experimenter in the lab, lack of anonymity between the subject and the experimenter, the fact that subtle (and potentially inadvertent) manipulations of context can have large effects on behaviour, low stakes compared to decisions outside the lab, and subject self-selection into the experiment. Some of these concerns apply mostly to experiments designed to observe a certain behaviour (say, positive giving in the dictator game) from which general behaviour is inferred (people are altruistic) and less to experiments featuring a within design, comparing with each other the choices made by different groups.

Two important remaining concerns are that the undergraduate student subjects we use in our laboratory experiment might be somehow different from the rest of the population and that people might behave differently in the lab than they would in other contexts. We will address these concerns using survey data on altruism and reciprocity outside of the lab covering a large and diverse population. The aim is to see whether the effects observed in the lab carry over to some of the behaviours the social preference games are intended to predict and to show that they apply to a diverse and international population.

We collected handedness data from responders to the LISS panel, an ongoing Internet panel covering 5000 households comprising 8000 individuals which were selected to be representative of the Dutch population. The participating individuals have already responded to a large number of questionnaires, some of which contain measures of altruism, trust and reciprocity. We asked them whether they are left-handed or right-handed but did not gather any additional information and our analysis therefore makes use of measures that are contained in the existing LISS datasets. In total, 5823 responders answered our handedness questions but the number of observations used in our regressions depends on the number of respondents who replied both to our questions and to the social preference questions already contained in the dataset. A potential worry with this more diverse sample is that amongst the older responders there might be left-handers who were forced to switch to using their right hand by their parents or teachers. This practice was common in the past but has become rare in western countries over the past decades (McManus, 2002). For this reason, we exclude subjects of sixty years or older from our sample which leaves us with 4073 observations.[12]

[12]We also asked responders how strong their preference for left or right was on a scale from 1 – not so strong – to 5 – very strong – and indeed we find that subjects who are sixty years of age or older are much more likely to indicate a weak preference for their hand of choice (Wilcoxon rank-sum test: $p \ll 0.01$). This phenomenon is much

The data available on (varying subsamples of) the panel members spans dozens of questionnaires containing thousands of variables. Apart from the handedness data we collected, we will limit ourselves to two questionnaires containing measures that are directly relevant to our research. The first study we use is the "European Social Survey" module which contains attitude questions concerning trust and reciprocity. The second study we use is the "Social Integration and Leisure" module of the LISS core study. This questionnaire contains data on altruistic behaviour such as donating money to charitable organisations, doing volunteer work, and performing informal care for another person.[13]

The 1997 wave of the National Longitudinal Survey of Youth (NLSY)[14], a US panel survey covering a nationally representative sample of individuals who were 12 to 16 years old at the end of 1996, contains both information on handedness and a limited number of indicators of altruism. The 2007 questionnaire contains binary indicators for donating money and doing volunteer work which are also available on the LISS panel so that we can further check the external validity of our findings by investigating whether they apply to the US population. All respondents were born between 1980 and 1984 and age-related issues do therefore not come into play.[15]

4.5 Survey results

4.5.1 Trust and reciprocity

Our laboratory results show that left-handers give substantially more in the trust and ultimatum games and that this effect is present in men but not in women. Given that left-handed men do not appear more altruistic in the dictator and public good games, this indicates that they trust others more and believe them to be ready to exercise negative reciprocity. It is hard to see what exactly would constitute a behavioural survey measure of trust and reciprocity, but the "European Social Survey" module of the LISS panel contains three attitude questions which cover different aspects of the respondent's beliefs about other people's reciprocity and trustworthiness, namely whether the respondent believes that others are fair, whether others can be trusted, and whether others deserve

more significant in right-handers (p<<0.01) than in left-handers (p=0.24), indicating that amongst the older subjects there are many individuals whose preference for right is a weak one. Also, there is no significant relationship between mixed-handedness (defined as indicating a preference strength of 3 or lower) and age for subjects under sixty years of age (OLS regression: p=0.34) while this relationship is highly significant for subjects of sixty years or older (p<0.01).

[13] A description of the two modules, the questionnaires, and the data can be found online at http://www.lissdata.nl/dataarchive/study_units/view/59 and at http://www.lissdata.nl/dataarchive/study_units/view/6.

[14] http://www.bls.gov/nls/y97summary.htm

[15] The panel oversamples the black and hispanic populations which is not an issue since handedness does not vary with race (p=0.593, Fisher's exact test).

Table 4.5: LISS trust sample

	Sample	Men	Women
N	1357	586	771
Left-handed	163	80	83
Age	42.3	42.0	42.6
Monthly income (€)	1451	2047	999
Education level	3.8	3.9	3.7

Education is divided into six categories according to the definition used by Statistics Netherlands.

to be trusted. These questions are "Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?" (Fairness); "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?" (Trust); and "Would you say that most people deserve your trust or that only very few deserve your trust?" (Trust-deserving). Respondents can answer each question on a scale from 0 to 10 whereby a higher number represents more trust.[16]

The Fairness and Trust questions have been used in many previous studies on trust and have been linked to choices in social preference games in the lab.[17] They have been found, for instance, to predict both giving and reciprocating in the trust game (Glaeser et al., 2000; Holm and Danielson, 2005). Fehr (2009), reviewing the trust literature, concludes that answers to trust related survey questions are driven by similar preferences and beliefs as decisions in the trust game.

Table 4.6 regresses the answers to these attitude questions on a left-handedness dummy and a set of controls. We can see that left-handed men, but not left-handed women, are indeed more likely to expect others to be fair, to think that most people can be trusted, and to find that most people are deserving of their trust. These results confirm our experimental finding of higher levels of trust in left-handed men and no effect for left-handed women.

4.5.2 Altruistic behaviour

In this section will concentrate on the analysis of the correlation between handedness and altruistic behaviour outside of the lab. Both the LISS and NLSY data sets do contain some straightforward

[16]For Fairness, 0 represents "Most people would try to take advantage of me" and 10 represents "Most people would try to be fair"; for Trust 0 represents "You can't be too careful" and 10 represents "Most people can be trusted"; and for Trust-deserving 0 represents "Very few people deserve my trust" while 10 represents "Most people deserve my trust".

[17]They have originally been taken from the American General Social Survey (GSS) which has used them to measure trust since 1972.

Table 4.6: Handedness, trust and reciprocity (LISS panel)

	(1)	(2)	(3)	(4)	(5)	(6)
			Fairness			
	Full Sample:		**Men:**		**Women:**	
Left-handed	0.142	0.224	0.407*	0.480**	-0.066	-0.005
	(0.142)	(0.142)	(0.227)	(0.225)	(0.176)	(0.173)
N	1357	1357	586	586	771	771
Mean	6.158	6.158	5.986	5.986	6.288	6.288
SD	1.774	1.774	1.864	1.864	1.692	1.692
			Trust			
Left-handed	0.132	0.215	0.492*	0.598**	-0.188	-0.110
	(0.184)	(0.181)	(0.268)	(0.274)	(0.251)	(0.234)
N	1357	1357	586	586	771	771
Mean	5.841	5.841	5.775	5.775	5.891	5.891
SD	2.168	2.168	2.223	2.223	2.126	2.126
			Trust-deserving			
Left-handed	0.052	0.113	0.409*	0.486**	-0.257	-0.203
	(0.152)	(0.150)	(0.243)	(0.237)	(0.183)	(0.179)
N	1357	1357	586	586	771	771
Mean	6.525	6.525	6.422	6.422	6.603	6.603
SD	1.794	1.794	1.898	1.898	1.709	1.709
Controls	no	yes	no	yes	no	yes
Scale	0-10	0-10	0-10	0-10	0-10	0-10

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; controls consist of age, gender, income and education

Table 4.7: LISS altruism sample

	Sample	Men	Women
N	3691	1620	2071
Left-handed	430	218	212
Age	40.3	40.4	40.2
Monthly income (€)	1372	1949	920
Education level	3.57	3.60	3.54

Table 4.8: NLSY sample characteristics

	Sample	Men	Women
N	7673	3862	3811
Left-handed	923	521	402
Age	24.9	24.9	25.0
Yearly income (US$)	20412	23528	17125
Non-white	3744	1851	1893

measures of altruism. Based on the experimental results from the lab discussed above, we would expect left-handedness to have a negative impact on altruism especially in women.

The "Social Integration and Leisure" module of the LISS core study contains several measures of altruistic behaviour, namely binary indicators for donating money, doing voluntary work, and performing informal care for a sick relative or friend. The data also contains information on membership in charitable organisations for "humanitarian aid, human rights, minorities or migrants". Membership in such an organisation arguably indicates a preoccupation with the common good and thus an altruistic streak. Table 4.9 reports regression results for the effect of handedness on these behaviours. We also show regressions controlling for gender age, gross monthly income and education as these can be expected to have an impact especially on donations.[18]

The donation regressions in Table 4.9 confirm our lab results on altruism: left-handers as a whole are significantly less likely to donate money and this effect is completely attributable to women. After inclusion of controls, female left-handers are around 5.5 percentage points less likely to donate money, which is a sizeable reduction compared to an average proportion of donors of around 38 percent. The picture is less clear for volunteering though. Left-handers as a whole are not

[18] The "Social Integration and Leisure" module of the LISS core study asks respondents whether they donated money to or volunteered for a range of organisations including humanitarian, environmental, and social. Our binary indicators are composite measures which indicate that an individual has donated to or volunteered for at least one type of organisation.

Table 4.9: Handedness and altruistic behaviour (LISS panel)

	(1)	(2)	(3)	(4)	(5)	(6)
	Donating money					
	Full Sample:		**Men:**		**Women:**	
Left-handed	-0.041*	-0.029	-0.007	-0.000	-0.069**	-0.055*
	(0.024)	(0.024)	(0.035)	(0.034)	(0.034)	(0.033)
N	3691	3691	1620	1620	2071	2071
Mean	0.376	0.376	0.354	0.354	0.392	0.392
SD	0.484	0.484	0.478	0.478	0.488	0.488
	Volunteering					
	Full Sample:		**Men:**		**Women:**	
Left-handed	-0.020	-0.012	-0.075**	-0.074**	0.039	0.047
	(0.024)	(0.024)	(0.031)	(0.032)	(0.035)	(0.035)
N	3692	3692	1621	1621	2071	2071
Mean	0.319	0.319	0.304	0.304	0.330	0.330
SD	0.466	0.466	0.460	0.460	0.470	0.470
	Informal care					
	Full Sample:		**Men:**		**Women:**	
Left-handed	-0.005	0.009	0.029	0.030	-0.019	-0.005
	(0.020)	(0.020)	(0.025)	(0.026)	(0.031)	(0.030)
N	3693	3693	1621	1621	2072	2072
Mean	0.195	0.195	0.122	0.122	0.253	0.253
SD	0.396	0.396	0.327	0.327	0.435	0.435
	Membership in humanitarian organisations					
	Full Sample:		**Men:**		**Women:**	
Left-handed	-0.018*	-0.014	-0.012	-0.011	-0.022	-0.017
	(0.009)	(0.009)	(0.012)	(0.012)	(0.014)	(0.014)
N	3691	3691	1619	1619	2072	2072
Mean	0.049	0.049	0.038	0.038	0.057	0.057
SD	0.215	0.215	0.190	0.190	0.233	0.233
Controls	no	yes	no	yes	no	yes

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; controls consist of age, gender, income and education.

Table 4.10: Handedness and altruistic behaviour (NLSY 97)

	(1)	(2)	(3)	(4)	(5)	(6)
	Donating money					
	Full Sample:		**Men:**		**Women:**	
Left-handed	-0.023	-0.014	0.006	0.007	-0.047**	-0.040*
	(0.014)	(0.014)	(0.018)	(0.018)	(0.022)	(0.022)
N	7673	7673	3862	3862	3811	3811
Mean	0.218	0.218	0.179	0.179	0.258	0.258
SD	0.413	0.413	0.383	0.383	0.438	0.438
	Volunteering					
	Full Sample:		**Men:**		**Women:**	
Left-handed	-0.037**	-0.031**	-0.031	-0.029	-0.037	-0.030
	(0.015)	(0.015)	(0.019)	(0.019)	(0.023)	(0.022)
N	7673	7673	3862	3862	3811	3811
Mean	0.256	0.256	0.234	0.234	0.279	0.279
SD	0.437	0.437	0.423	0.423	0.449	0.449
Controls	no	yes	no	yes	no	yes

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; controls consist of age, gender, income and race

significantly less keen on doing volunteer work. But the effect is strongly and significantly negative for left-handed men – a reduction of 7.4 percentage points relative to an average propensity of 32 percent – while the coefficient is insignificant and positive for left-handed women. This gender difference is rather unexpected given the lab findings. Left-handers are neither more nor less likely to perform informal care for others. Finally, we find that left-handers are less likely to be a member of a humanitarian organisation. This effect is present for both men and women separately but is only significant at the 10-percent level for the sample as a whole (economically, at 1.8 percentage points the effect is relevant relative to a sample average of 4.9 percent).

Table 4.10 contains regression results for the NLSY sample. These results, based on a sample of American twenty-somethings, mostly confirm the results obtained from the Dutch LISS sample. Again, left-handed women are significantly less likely to have donated money while the effect for men is zero. The magnitude of the effect is 4.0 percentage points which is sizeable compared to a sample average of 21.8 percent. We also find that both men and women are around 3 percentage points less likely to volunteer although the effect is only significant for the sample as a whole. Here the NLSY sample differs somewhat from the LISS sample where a negative effect of left-handedness on volunteering was only observed for the male sub-sample.

Donations are arguably the closest equivalent to the dictator game in as far as the final recipient usually does not know the identity of the donor and has no opportunity to reciprocate. The donation regressions therefore confirm our lab results on altruism as they replicate the pattern found in the lab for the dictator game both with the LISS data and the NLSY data: left-handed women are significantly less likely to donate while there is no significant difference for men. The survey data results nevertheless also confirm that one needs to be cautious when generalising lab results on social preferences to predict pro-social behaviour in general. Volunteering is a very different activity from donating money as it is much more visible and requires more personal involvement. But it is still important to note that the lab results from the dictator and public good games do not indicate the negative correlation between left-handedness and volunteering for males. This finding illustrates that contextual factors, which are eliminated in the lab, can indeed lead to different effects in the field.

4.6 Possible biological mechanisms

Handedness is related to a range of neurological differences and identifying the exact mechanism by which handedness affects social preferences is therefore difficult. In this section, we nevertheless discuss some potential mechanisms, an endeavour which necessarily remains speculative.

As discussed in the introduction, male left-handers have a thicker corpus callosum. This results in an increased capacity for communication between the two hemispheres and suggests that the brains of left-handed men are better integrated in processing information.[19] For instance, left-handers perform better at activities that require rapid transfer of information, such as communication and empathising (Hines et al., 1992).[20] This potentially indicates higher levels of empathy in left-handers, especially for men.

The strength of prenatal exposure to testosterone may be both positively or negatively correlated with left-handedness (Grimshaw et al., 1995). Prenatal testosterone exposure is thought to shift brain development away from a brain geared towards empathy in direction of a brain wired for dealing with systems.[21] It is also a potential cause of autism, a condition associated with very low levels of empathy (Manning et al., 2001). More recent studies point towards a negative correlation of prenatal testosterone exposure with left-handedness especially for men (Lust et al., 2010), implying higher levels of empathy in male left-handers. A positive correlation, implying lower levels

[19] Witelson (1985); Witelson and Goldsmith (1991); Grimshaw et al. (1995)

[20] Also see Baron-Cohen (2003) and the references therein.

[21] See Baron-Cohen et al. (2004) for a review of the large body of research backing the testosterone-empathy link and Baron-Cohen (2003) for a more accessible account.

of empathy in left-handers, seems more likely for women.[22]

Increased empathy has been shown to lead to higher levels of giving in the ultimatum game (Barraza and Zak, 2009) and it is thought that hormones influence altruism and trust through an impact on empathy (Zak, 2011). A link between handedness and empathy via the two above-mentioned neurological mechanisms is therefore potentially behind our results.

4.7 Conclusions

This paper contributes to our knowledge about the determinants of individual differences in social preferences. People differ strongly from each other in their degrees of altruism, trust and reciprocity and the fact that these differences occur even in the tightly controlled environment of lab experiments suggests that, rather than being determined on the spot by circumstantial factors alone, they may be determined by deep-seated differences in preferences. What determines these preferences, we have only just begun to explore. We present evidence from a controlled lab experiment and two large online surveys showing that handedness is significantly correlated to individual levels of altruism, trust and reciprocity. Handedness is a well-documented predictor of neural differences and our results consequently hint at a neural basis for social preferences.

Our lab results show that left-handed men make choices that are markedly different from those of right-handed men when the affected individuals have a chance to reciprocate. Left-handed men are also themselves more likely to exhibit positive reciprocity. These effects are not present in left-handed women who, on the other hand, are less altruistic in the dictator game where reciprocity does not play a role. A potential explanation for the behaviour of male left-handers is that they have a stronger belief in the trustworthiness and reciprocity of others. This would mean that they give more not out of a desire to make the responder better off but because they believe, in the case of the trust game, that their trust will be rewarded and, in the case of the ultimatum game, that low offers would be rejected.

Connecting our lab results to large-scale survey data, we can demonstrate the external validity of our findings to a degree which is usually difficult to accomplish for studies investigating social preferences. The lab result of lower altruism in left-handers translates to lower rates of charitable giving, lower willingness to do volunteer work, and lower rates of membership in humanitarian organisations. Especially for donations, arguably the closest equivalent of the dictator game, the observed patterns in both the Dutch and the US data are consistent with the patterns found in the

[22]Using the same sample as this study, Buser (2012) finds that physical markers for the strength of prenatal testosterone exposure predict lower giving rates in both men and women.

lab. We also use a series of attitude questions which gauge the respondent's beliefs about the fairness and trustworthiness of others to show that left-handed men, but not women, are indeed more likely to think that others are trustworthy and fair, again confirming our lab results. But it needs to be noted that we observe gender differences for some of the survey measures of altruism which do not correspond to those found in the lab. We therefore conclude that one nevertheless needs to be careful when generalising experimental results on social preferences to predict pro-social behaviour in general.

Left-handedness, particularly in men, is associated with a stronger integration of the two hemispheres of the brain and therefore of brain functions. A tentative explanation for our findings is that these differences in neural structures lead to differences in reciprocity and beliefs about other people's reciprocity, potentially because male left-handers find it easier to process and take into account information about others and reckon with their possible intentions. More specifically, our results indicate that brain lateralisation and the size of the corpus callosum play an important role in shaping social preferences.

In summary, our contribution is twofold. Firstly, we show that the neural differences associated with handedness lead to different choices in social preference games in the lab. The observed differences are particularly strong for reciprocal behaviour, which is inextricably linked to our expectations and beliefs concerning the behaviour and motivations of others. Secondly, we examine the link between biological factors and social preferences in the field as well as the lab. We can show that our findings partially carry over to the field where we observe strong effects on the prevalence of altruistic behaviour and trust in large and representative samples from the Netherlands and the US. We conclude that social preferences are at least partially biologically determined.

Chapter 5

Gender, Competition and Career Choices[1]

5.1 Introduction

Gender differences in labor market outcomes, while greatly reduced, have remained ubiquitous. One driving source for the gender wage gap seems to be gender differences in education. To understand these gender differences in career and educational choices, psychological and socio-psychological attributes are now commonly discussed as potential explanations. While the last decade saw a flurry of laboratory evidence on gender differences on psychological attributes (see Croson and Gneezy, 2009), the direct evidence linking the experimental literature to outcomes in the education and labor market has been rather scant.[2] This paper aims to close that gap and contributes to both these literatures. Specifically, we investigate to what extent one of the most robust gender differences in laboratory experiments, the gender difference in competitive attitudes, can help account for gender differences in educational choices. We do this by measuring competitiveness among school children for whom we also have educational outcomes such as grades, and linking this data to subsequent educational choices.

Though gender differences in educational choices are smaller than they used to be, they remain significant. While in the U.S. girls take on average as many advanced math and science classes as boys and perform on average at similar levels (Goldin et al., 2006) this is not the case in all OECD countries.[3] Even in the U.S., girls are underrepresented among extremely high achieving

[1]The research in this chapter is joint work with Muriel Niederle and Hessel Oosterbeek.

[2]Bertrand (2011) summarizes this literature and concludes: "While the laboratory evidence shows in many cases large gender differences (say, in attitudes towards risk, or attitudes toward competition), most of the existing attempts to measure the impact of these factors on actual outcomes fail to find large effects. This is undoubtedly a reflection of a rather new research agenda, as well as of the difficulty in finding databases that combine good measures of psychological attributes with real outcomes."

[3]We will show that in the Netherlands boys are significantly more likely to take math classes in high school

math students. Ellison and Swanson (2010) provide compelling evidence that this gender gap is not driven solely by differences in mathematical ability. Specifically, they show that in mathematics, high-achieving boys come from a variety of backgrounds, while high-achieving girls are almost all drawn from a small set of super-elite schools. At the college level, even in the U.S. women are significantly less likely to graduate from a so-called STEM (science, technology, engineering and mathematics) major than men.[4] Research investigating career choices of women and men suggests that among equally gifted students, males are much more likely to choose a math heavy college major (see LeFevre et al., 1992; Weinberger, 2005).

The reason to be concerned about gender differences in math and sciences, compared to, say, literature, is that the choices of math and science classes are most predictive of college attendance and completion (Goldin et al., 2006). Furthermore, performance in mathematics has consistently been found to serve as a predictor for future earnings. For example, Paglin and Rufolo (1990) report that a large fraction of the gender gap in average starting salaries for college graduates is between rather than within detailed college majors (for additional evidence and discussion see Grogger and Eide, 1995; Brown and Corcoran, 1997; Weinberger, 1999; Weinberger, 2001; Murnane et al., 2000; Altonji and Blank, 1999).[5]

To assess the effects of psychological attributes on educational choices, we want to measure them before students make choices which result in different educational experiences. This ensures that measured psychological attributes are not contaminated by the different environments children experience after having selected into different educational careers. We run our study in the Netherlands where, at the end of the third year of secondary school, students in the pre-university track have to choose between four study profiles: a science-oriented profile, a health-oriented profile, a social science-oriented profile and a humanities-oriented profile.[6] There is a clear ranking of these profiles in terms of academic prestige, with the science-oriented profile being the most and the humanities-oriented profile being the least prestigious and challenging. Despite the fact that girls perform slightly better academically, boys enroll disproportionally often into the science-oriented profile, while girls enroll disproportionally often into the health-oriented and humanities-oriented profiles. The choice of study profile in secondary school is furthermore strongly correlated with the choice of major in tertiary education which in turn is strongly correlated with future occupation

than girls. In France, where like in the Netherlands high school children decide on which set of classes to enroll in, girls are less likely to choose the math and science heavy options (http://media.enseignementsup-recherche.gouv.fr/file/2010/42/2/filles-garcons-egalite-ecole-a-enseignement-superieur2010_139422.pdf).

[4] http://nces.ed.gov/pubs2009/2009161.pdf

[5] In a study on the gender gap in earnings among MBA's from Chicago Booth, Bertrand et al. (2011) conclude that one of three factors that account for the large gender gap in earnings a decade after MBA completion is differences in training prior to MBA graduation, with, most notably, women taking many fewer finance courses than men.

[6] In the Netherlands, students are selected into tracks at the end of primary school, at age 12, with about 20 percent of the highest performing students enrolling in the pre-university track during high school.

and therefore with future labor market positions and earnings.

We aim to assess to what extent choices of study profiles are driven by psychological attributes, most notably competitiveness, as well as grades and students' beliefs about their own talent at math. We focus on competitiveness because gender difference in competitive attitudes is the largest and most robust gender difference found in experimental studies (see Gneezy et al., 2003, Niederle and Vesterlund (2007) and, for an overview, Niederle and Vesterlund, 2011). In the first paper on gender differences in preferences for performing under competitive or non-competitive incentive schemes, Niederle and Vesterlund (2007) assess choices of college students that perform equally well in a simple arithmetic task. They find that while 73 percent of men choose a competitive tournament payment scheme instead of non-competitive piece-rate compensation, only 35 percent of women do so. Subsequent research has confirmed this gender difference in the willingness to compete in somewhat stereotypical male tasks, such as simple arithmetic problems (see Niederle and Vesterlund, 2011). Sutter and Rützler (2010) provide evidence that these gender differences in competitiveness are present among children of all ages, beginning when they are 3 years old.

Niederle and Vesterlund (2007) show that gender differences in competitiveness are largely due to gender differences in confidence and competitive attitudes, with gender differences in risk aversion playing a minor role. We therefore aim to decompose the effect of competitiveness on educational choices into pure competitive attitudes, as well as confidence and risk, which are other well-studied traits that show significant gender differences (see Croson and Gneezy, 2009).

The fact that women shy away from competition while men compete too much has potentially important implications for labor market outcomes. People who shy away from competitive environments may self-select into different, potentially lower paid, careers. Gender differences in competitive attitudes may therefore be a potential explanation for the under-representation of women in certain fields such as the sciences which are viewed as competitive.[7]

The goal of this paper is to assess the extent to which gender differences in competitiveness can account for gender differences in educational career choices. Among a set of four schools in and around Amsterdam, we administered an experiment just prior to the moment when students make their first important educational choice. We elicit the students' competitiveness using the design of Niederle and Vesterlund (2007), as well as their confidence and risk aversion. We also assess the students' beliefs about their mathematical prowess, as grades may not be the most accurate predictor of mathematical ability. The school provided us with the subsequent profile choices of

[7] The fact that women shy away from competition more than men may also account for the fact that few qualified women reach the top. Women only account for 2.5 percent of the five highest paid executives among a large dataset of U.S. firms (see Bertrand and Hallock, 2001). Furthermore, it may account for why the gender log wage gap accelerates in the upper tail, as documented by Albrecht et al. (2003) for Sweden, and later confirmed in other studies Arulampalam et al. (e.g. 2007).

students as well as their grades.

As expected, we find that boys are more than twice as likely than girls to compete. We also find that competitiveness is strongly related to profile choice. Competitive students choose more prestigious profiles. This finding is robust to the inclusion of control variables, including grades, self-rated ability, (over)confidence and risk tolerance. Using ordered probit estimation, we find that our measure of competitiveness can explain around 20 percent of the gender difference in profile choice.

The remainder of this paper is organized as follows. Section 5.2 provides details of the structure of Dutch secondary education and of the study profiles. Section 5.3 discusses the design of our study and describes the data. Section 5.4 presents and discusses the results. Section 5.5 summarizes and concludes.

5.2 Academic study profiles in the Netherlands

The students participating in this study are drawn from the population of Dutch secondary school students who are enrolled in the pre-university track. In the Dutch school system, tracking takes place when students go from primary school - grade 1 to 6 - to secondary school, normally at age 12. There are three tracks: a six-year pre-university track, a five-year general track and a four-year vocational track.[8] Around 20 percent of students graduate from the pre-university track, 25 percent from the general track and the remaining 55 percent from the vocational track. Who enrolls in which track is to a large extent determined by the score on a nation-wide achievement test administered at the end of primary school. This test consists of multiple choice questions dealing with language, arithmetic/mathematics, information processing and (optionally) world orientation. It supposedly measures students' cognitive ability. Our sampling frame consists of the 20 percent pre-university track students.[9]

Halfway through the six years of secondary school, at the end of grade 9, students in the pre-university track have to choose between four study profiles: the science-oriented profile Nature & Technology (NT), the health-oriented profile Nature & Health (NH), the social science-oriented profile Economics & Society (ES) and the humanities-oriented profile Culture & Society (CS). Table 5.1 shows the subjects that differ across the different study profiles and the number of teaching hours assigned to these subjects in the last three years of secondary school.[10] Mathematics is

[8]The latter is divided into different sub-tracks which differ in the shares of school-based and work-based learning.

[9]Girls are somewhat more likely than boys to go to the pre-university track, making up 54 percent of the students, (source: Statistics Netherlands (CBS)).

[10]In addition the students take the following non-profile specific subjects: Dutch (480 hours), English (400), second

Table 5.1: Subjects and teaching hours per study profile

Nature & Technology - NT	**Nature & Health - NH**
Mathematics B - 600	Mathematics A - 520
Physics - 480	Biology - 480
Chemistry - 440	Chemistry - 440
Nature, life and technology – 440	Nature, life and technology – 440
or IT - 440	or geography - 440
or biology - 480	or physics - 480
or mathematics D - 440	
Economics & Society - ES	**Culture & Society - CS**
Mathematics A - 520	Mathematics A or C - 480
Economics - 480	History - 480
History - 440	Art – 480
Management and organization – 440	or philosophy – 480
or geography – 440	or modern foreign language - 480
or social studies - 440	or Greek or Latin - 600
or modern foreign language - 480	Geography – 440
	or social studies - 440
	or economics - 480

Note: The table lists the subjects per profile and the number of teaching hours per subject during the last three years of the pre-university track. Source: Ministry of Education, Culture and Science.

taught in each track, but at different levels; A, B, C and D, where D is the most advanced followed by B, A and C.

Some schools also allow for combined profiles, namely NT/NH and ES/CS. In the NT/NH profile, students take the hardest mathematics version, Mathematics B, albeit only at 520 hours. Furthermore, Physics is not required. In the ES/CS profile, students replace one of the CS-electives with the economics course. As such, the combined profiles are in between the pure profiles, though a little closer to NT and ES, respectively.

The choice of study profile in secondary school is strongly correlated with the choice of major in tertiary education.[11] Table 5.2 shows for each study profile the distribution of students across undergraduate majors. Most NT graduates study a subject in science and engineering, NH graduates often opt for health-related subjects, ES graduates often choose a major in economics and business or in law, and most CS graduates choose a subject in the humanities, social sciences or law.[12] Different study profiles are not only associated with different careers, they also differ in the

foreign language, Latin or Greek (480), social studies (120), general natural sciences (120), culture (160), and sports (160). The students spend roughly half their time on profile specific subjects and half on common subjects.

[11] Undergraduate systems in European countries are different from the US. In the US, people start by sampling lots of courses, then decide on a major later. In Europe, students choose a major from the beginning of their studies and only take major-relevant courses.

[12] Some studies actually restrict entry to students who took certain profiles or courses within profiles. For example,

Table 5.2: Undergraduate major by profile (percentages)

	NT	NH	ES	CS
Humanities	9	6	8	30
Social Sciences	2	9	19	34
Law	1	4	20	20
Economics and Business	15	8	46	5
Science and Engineering	64	18	2	0
Health Care	7	48	1	1
Other	2	7	4	9
Going to university	81	72	69	60
Profile Choices				
Boys	35	21	38	6
Girls	10	34	32	24

Source: Statistics Netherlands (CBS). The data from the top rows are from 2006. The data from the bottom rows are from 2009, where we exclude choices of combined profiles.

likelihood with which students in each profile enroll in university. While 81 percent of NT students continue their education at the university level, only 60 percent of CS students do so. This ordering of study profiles also corresponds to how the profiles are viewed. NT is generally regarded as the most challenging and highest-reward study profile, followed by NH and ES, and CS as the least demanding and lowest-reward study profile.

In year three of the pre-university track, at the end of which students decide upon their study profile, girls perform overall somewhat better than boys. They are, for instance, less likely to drop out or repeat a year. In standardised tests such as the PISA test[13], girls and boys perform similarly whereby boys do slightly better at math and girls slightly better at other subjects (Driessen and Van Langen, 2010).

Despite these similarities between boys and girls, they make very different study profile choices. Table 3 shows that boys are more likely to choose more prestigious study profiles. Compared to girls, boys are more than three times as likely to choose the most prestigious profile, NT, and less than a third as likely to choose the least prestigious profile, CS. The fact that girls are disproportionately more likely to choose CS has prompted a debate with the minister for education even proposing to eliminate the profile altogether. This idea was ultimately rejected and the profiles are to stay as they are for a while to come.[14]

medical schools require NT or NH; to study math, having taken Math B in high school is required.

[13]Programme for International Student Assessment; an evaluation in OECD member countries of 15-year-old school students' academic performance.

[14]Source: http://nos.nl/artikel/203421-minister-wil-onderwijs-reorganiseren.html and http://nos.nl/artikel/268284-raad-niet-minder-profielen-havovwo.html

5.3 Experimental subjects and design

The aim of this paper is to assess the extent to which psychological attributes, most notably competitiveness, can account for gender differences in academic profile choice in the Netherlands. This can serve as a test to show that psychological attributes can account for gender differences in labor market outcomes. To achieve this, we focus on a small set of students in the Netherlands who are about to choose their study profile. We assess their psychological attributes through a classroom experiment. In addition we collected information through a short questionnaire and received administrative data provided by the schools. This will allow us to study the effect of psychological attributes after controlling for classic channels that may account for profile choices, such as the students' grades. In this section we describe the environment and the participants in our study, followed by the experimental design for assessing the psychological traits.

5.3.1 The environment

We invited secondary schools in and around Amsterdam to participate in a research project investigating the determinants of academic profile choices. We demanded one class hour (45 or 50 minutes) of all classes in 9th grade, the 3rd grade in the pre-university track. The invitation letter stated that students would participate in an experiment and be paid depending on their choices.

Four schools cooperated, one in the city of Amsterdam and three in cities close to Amsterdam. In each school, we captured all students in the 3rd grade of the pre-university track, that is students could not self-select into the experiment. The number of classes per school varied from 3 to 5; in total we have data from 16 classes. A total of 397 students participated in the experiment.

After the end of the school year, the schools provided us with the final grades, including mathematics, and the definite profile choices of each student. For 35 students we do not have a definite profile choice.[15] For 20 of these students, we use the profile choice from the questionnaire.[16] We drop the remaining 15 students for whom we have neither a final choice nor a clear choice from the questionnaire. We have to drop an additional 4 students from the analysis because they showed up late to class and missed part of the experiment, 2 students because their questionnaires were incomplete and they therefore lack key control variables, and 14 students because we did not obtain their grades. This leaves us with a sample of 362 subjects.

[15] Some students may have to retake the year, and in some schools those are included in the final profile choice, in some not.

[16] For the students for whom we have both the final profile choice and the intention stated in the questionnaire, the questionnaire answer accurately predicts the final choice in 94 percent of the cases.

5.3.2 The experimental design

We use a classroom experiment to obtain an individual measure of competitiveness. The design closely follows Niederle and Vesterlund (2007). Participants perform a real task, first under a non-competitive piece rate scheme and then under a competitive tournament scheme. Participants choose which of the two payment schemes to apply to their third and final performance. This allows us to determine the extent to which the choice of compensation scheme depends on performance.

The task of the experiment is to add up sets of four two-digit numbers for three minutes. The performance in each round corresponds to the number of correctly solved problems. In each round, participants received envelopes that contained a sheet of 26 problems. After having read out the instructions that were on top of the envelopes and answering potential questions, the experimenter gave the signal that subjects could open the envelopes and start the addition problems. Participants were not allowed to use calculators, but could use scratch paper. At the end of the three minutes, subjects had to drop the pen and stand up. In each round there were three versions of the 26 addition problems to prevent copying from neighbors.

Participants were informed they would perform in three rounds, one of which was randomly chosen for payment at the end of the experiment through the roll of a die in front of the class. Participants received details on each round only immediately before performing in the task. Students did not receive any information on the performance of anyone else and were paid, a week later, through sealed envelopes, at which time they could make inferences about their relative performance. Participants earned on average €5.55, with a minimum of zero and a maximum of €25.

Participants first performed the task under a noncompetitive piece rate of 25 Euro-cents per correctly solved problem. In round two they performed in tournaments of four, where the three competitors were randomly selected among students from the same class by computer after the end of the experiment. The person with the largest number of correctly solved problems would be paid €1 per correct problem and the others received no payment. In case of a tie, the winner was randomly determined. Participants did not receive any information about their own performance or the performance of others, including whether they won or lost the tournament, since their competitors where only determined after the experiment.

In the third round, participants chose between the two payment schemes. In case round three was selected for payment, the earnings were computed as follows. A participant who chose the piece rate received 25 cents per correct problem. A participant who selected the tournament would win if his or her new round 3 performance exceeded the performance of the other three group members in the previous round 2 tournament. Therefore, just like in Niederle and Vesterlund (2007), the choice was an individual decision as a participant's choice did not affect the payoffs of any other participant.[17]

We elicit two additional psychological attributes that may affect the decision to enter the tournament as well as the decision to choose a prestigious study profile in high school. The first is confidence, measured by the beliefs of participants about their relative performance. Gender differences in beliefs about relative performance, especially in performances in tournaments, are well documented and have been shown to explain a portion of the gender gap in tournament entry (see Niederle and Vesterlund (2007) and also Mobius et al., 2011). Gender differences in confidence may also influence the profile choices of students. We therefore elicited the students' beliefs about their relative performance after round 3. Specifically, we asked students about their relative performance in the round 2 tournament compared to the other three group members. If their guess was correct, they received €1.[18]

The final factor that may influence the choice of compensation scheme and academic careers is attitudes towards risk. While risk attitudes are correlated with tournament entry (see for example Buser, 2011), note that the literature so far has failed to find that risk attitudes play a prominent role in accounting for gender differences in the choice of compensation schemes (see Niederle and Vesterlund, 2011). We elicited risk attitudes by using two measures. First, following Eckel and Grossman (2002), subjects picked one option among a sure payoff of €2 and four 50/50 lotteries with linearly increasing riskiness and expected payoffs: 3 or 1.5; 4 or 1; 5 or 0.5; 6 or 0. The outcome of the lottery was determined by a dice roll at the end of the experiment.

Second, we asked subjects "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" The answer is on a scale from 0 ("unwilling to take risks") to 10 ("fully prepared to take risk"). This second risk measure has advantages and disadvantages. First, it is simply a survey question, which makes it cheap, but potentially less reliable. Dohmen et al. (2011), using representative survey data from Germany, find that this simple question predicts both choices in a lottery task and risky behavior across a number of contexts including holding stocks, being self-employed, participating in sports, and smoking. Lonnqvist et al. (2010) find the question to be more stable over time than lottery measures for risk attitudes.

The experiment concluded with a questionnaire. Mathematical ability can be expected to be an

[17]There are several advantages to having participants compete in round 3 against the previous round 2 tournament performance. First, the performance of a player who chose the tournament is evaluated against the performance of other players in a tournament. Second, the choice of compensation scheme of a player does not depend on the choices of other players. Third, the participant provides no externality to another subject, hence motives such as altruism, or fear of interfering with someone else's payoff play no role.

[18]When two subjects have the same number of correctly solved additions they receive the same rank. For example, if two subjects are tied for first place, they are both ranked first and receive €1 if their guessed rank is equal to 1. The next best subject is ranked third.

important factor when choosing academic profiles. It may well be that grades are not a perfect predictor of mathematical ability. We therefore wanted additional measures of mathematical prowess. Specifically, we asked students to rank themselves on mathematical talent compared to other students in their year (and school) on a scale from 1 (the best 25%) to 4 (the worst 25%).[19] We also asked students how difficult they find it to pass their math class on a scale from 0 (very easy) to 10 (very hard).[20] Finally, we asked students which profile they expected to choose in June, several weeks after the experiment. The experiments were conducted in March, April and May of 2011.

5.4 Results

We present the results in three stages. First, we describe the study profile decisions of participants and the effect of grades on gender differences in choices. We document significant gender differences in profile choices that are not explained by grades or beliefs about mathematical prowess. Second, we assess gender differences in competitiveness. We analyze the extent to which other psychological attributes such as confidence and risk attitudes can account for gender differences in competitiveness. We find a significant gender difference in competitiveness. About 25 percent of this gap can be attributed to gender differences in confidence, while risk aversion can explain an additional 10-15 percent, leaving a large part of gender differences in tournament entry unaccounted for. In the main results section we examine to what extent gender differences in competitiveness can account for gender differences in study profile choices.

The competitiveness measure from the experiment significantly affects profile choice even conditional on real and perceived ability. The key finding is that gender differences in competitiveness can account for around 20 percent of gender differences in study profile choices. In the last part we decompose the effect of competitiveness on study profile choices into the effects of competitive attitudes, confidence and risk attitudes. The most important and robust psychological attribute is competitiveness.

5.4.1 Profile choices and school data

We first describe the profile choices as well as the grades and questions relating to mathematical prowess of the 362 students in our sample. Two of the four schools in our sample allow students

[19] This was phrased as three yes/no questions: "Do you think your mathematics ability is in the top 25% of your year?", "...top 50% of your year?", "...top 75% of your year?". A student who answered all 3 questions with a no was automatically assumed to be in the bottom 25%. We had 44 students who answered no to all questions. A student who answers yes to one of the questions also should answer yes to the next (if one is in the top 25%, one is also in the top 50%). 67 students, however, switched back to no. For these students, we count the first yes as their true answer.

[20] "How difficult is it for you to pass the math class?" (on a scale from 0 - very easy - to 10 - very hard).

Table 5.3: Descriptive statistics about profiles

	NT	NH	ES	CS	
All: Prestige (% rank)	1.48 (71%)	2.13 (57%)	2.64 (60%)	3.67 (81%)	
Boys: Prestige	1.43 (75%)	2.24 (57%)	2.59 (56%)	3.68 (82%)	
Girls: Prestige	1.52 (68%)	2.03 (57%)	2.71 (64%)	3.66 (80%)	
By chosen profile					Difference
GPA (1-10)	7.12	7.11	6.63	6.51	0.00
Math grades (1-10)	7.25	6.73	6.20	6.21	0.00
Math difficulty (0-10)	1.95	3.62	4.90	5.30	0.00
Math quartile (1(best)-4)	1.52	1.98	2.50	2.67	0.00
Observations	102	89	128	43	

Top rows: Average ranking of study profiles, and in parentheses, the fraction of students who rank that profile first (for NT), second, third or fourth (for NH, ES and CS, respectively). Bottom rows: Average characteristics of subjects who chose that profile. Grades are out of 10 with higher numbers being better grades. Math difficulty goes from 0 - very easy to 10 - very hard. Math quartile goes from 1 - best 25% to 4 - worst 25%. The last column reports p-values from Kruskal-Wallis tests.

to pick the combined profiles. Of the 173 students in those two schools, 64 students choose the NT/NH combination and 18 pick the CS/ES combination. For the main analysis of this paper we use the chosen profile as stated in the questionnaire to split the NH/NT and ES/CS combination choices into NH and NT and ES and CS decisions, respectively. However, since one can argue that the NT/NH profile is closer to NT, and the ES/CS closer to ES, we reestimate all regressions using this alternative definition of profile choice. Lastly, as a further robustness check, in the appendix, we show results where we treat NT/NH and ES/CS as separate categories.[21] The results remain qualitatively the same.

Before showing the profile choices of students, we confirm that the students in our sample rank the four profiles in the predicted order. We asked the 362 secondary school students in our sample to rank the four study profiles by asking "Which profile do the best students pick?". Their responses concur with the general opinion. The first row of Table 5.3 shows that the average rank follows the expected pattern, with NT having the best average rank, followed by NH and ES, and CS having the worst average rank. Furthermore, a majority of over 70 percent of students believes NT is the most demanding study profile. A majority of students ranks the NH study profile second and ES third. Finally, more than 80 percent of the students rank CS as the least demanding profile. As Table 5.3 shows, the rankings of boys and girls are remarkably similar. We also asked students to rank the four study profiles in terms of future earnings.[22] The picture that emerges is very similar;

[21]For these analyses we have to drop an additional 20 students for whom we do not have a final profile choice. The multiple choice question in the questionnaire did not allow for combined profiles and we do therefore not know whether these students would have picked one of the mixed profiles.

[22]The exact question was "With which profile do you think you would earn most in ten year's time? Rank the profiles from 1 to 4 where 1 means that you would earn most if you chose that profile and 4 that you would earn least

Table 5.4: Descriptive statistics by gender

	Boys	Girls	Gender-dif.
GPA (1-10)	6.76	6.97	0.01
Math grades (1-10)	6.67	6.59	0.64
Math difficulty (0-10)	3.41	4.18	0.01
Math quartile (1-4)	1.97	2.25	0.00
Profile Choice			
Nature & Technology (NT)	.40	.17	
Nature & Health (NH)	.12	.36	
Economics & Society (ES)	.39	.32	
Culture & Society (CS)	.08	.15	0.00
Observations	177	185	

Grades are out of 10 with higher numbers being better grades. Math difficulty goes from 0 - very easy to 10 - very hard. Math quartile goes from 1 - best 25% to 4 - worst 25%. The last column of the top four rows reports p-values from Wilcoxon ranksum tests and for the profile choice we report the p-value from a Fisher's exact test.

for NT the mode is the first rank, for NH the mode is the second place, for ES the mode is the third place and few students disagree that CS is the study profile with the poorest earnings prospects.[23]

The second part of Table 5.3 shows descriptive statistics by study profile. The choices conform with the view that higher performing children are more likely to choose more prestigious profiles. This is true both for the overall GPA, which is computed as the average of all grades, including mathematics, as well as for just the mathematics grade. Likewise, students that chose more prestigious profiles find mathematics less difficult (0 - very easy to 10 - very hard) and believe they are more likely to be among the highest ability children in mathematics (1 - best 25% to 4 - worst 25%).

Given that higher performing children choose more prestigious study profiles, we now assess academic differences between boys and girls in our sample in Table 5.4. On average, girls have a higher GPA than boys. In mathematics, there are no gender differences in performance. Despite this fact, girls say that they find it more difficult to pass the mathematics class and they rank themselves as less likely to be among the high math ability children in their school. This could reflect that girls are truly less able than boys in mathematics, as grades may be a function of true ability, as well as, for example, conscientiousness. On the other hand, the measure on self-assessed mathematical prowess may already include gender differences in confidence, and as such represent already a psychological attribute.

While academically boys and girls are very comparable, girls make significantly different profile

if you chose that profile." This question was only asked to students in two of the four schools and the percentages are therefore based on 181 observations.

[23] 50% think that NT gives the best salary prospects, 27% think NH, 20% ES and 2% CS.

Table 5.5: Gender and profile choice

	(1)	(2)	(3)
Female	0.342***	0.449***	0.329***
	(0.114)	(0.121)	(0.123)
Math Grade		-0.188	-0.016
		(0.133)	(0.141)
GPA		-0.360***	-0.326***
		(0.105)	(0.109)
Rel. Math Gr.		0.509	0.345
		(0.475)	(0.484)
Math Difficulty			0.336***
			(0.075)
Math Quartile			0.076**
			(0.031)
Cut 1 (C1)	-0.404	-3.957	-1.766
Cut 2 (C2)	0.251	-3.205	-0.948
Cut 3 (C3)	1.367	-1.961	0.392
Female/(C3-C1)	0.193	0.225	0.152
N	362	362	362

Dependent variable: Profile choice, where NT<NH<ES<CS. Coefficients from ordered probit regressions; standard errors in parentheses; *, ** and *** denote significance at 10, 5 and 1 percent, respectively.

choices compared to boys. The lower part of Table 5.4 shows profile choices by gender in our sample of 362 students. The pattern is similar to the pattern observed in national statistics. The NT profile is much more popular among boys than girls, while the opposite holds for NH. The ES profile is slightly more popular among boys than girls, and girls are more likely than boys to choose the least prestigious profile, CS. Note that in our sample, boys and girls are as likely to choose one of the Nature profiles compared to one of the Society profiles. This is not the case in national statistics, where girls are overall more likely to choose a Society profile, while boys are overall more likely to choose a Nature profile.

To more precisely understand gender differences in profile choice, we show in Table 5.5 ordered probit regressions where we order profiles by NT<NH<ES<CS. The first column shows that girls on average choose a significantly less prestigious profile. Specifically, the effect of being female bridges almost 20 percent of the distance between the best and the worst profile. We then include academic variables such as overall GPA, the mathematics grade and the mathematics grade compared to other children in the same class.[24] When we include these three academic achievement

[24]To compute the relative mathematics rank in class, we include all 397 students in our sample, including the 35 students we had to drop for the final results. Specifically, we gave the best students in class a rank of 1. The rank of each student is equal to 1 plus the number of students with a better grade. We then normalize the measure by dividing by the number of students in the class.

variables, the gender gap actually increases, and being female bridges almost 23 percent of the distance between the best and the worst profile. Note that the coefficient on female is actually slightly larger than the coefficient on the GPA. An increase in 1 in the GPA, which corresponds to about 144 percent of a standard deviation increase, corresponds to bridging 18 percent of the gap between the best and the worst profile. When we add the students' beliefs about their underlying mathematics ability and their belief about how good they are at math compared to their peers, the gender gap shrinks but remains large and highly significant. While these additional variables may already be viewed as psychological attributes, it may well be that they produce an additional insight into a students' real mathematical ability compared to grades only. In any case, there is a significant gender difference in study profile choice, with girls choosing less prestigious profiles than boys.[25]

Table 5.10 in the appendix shows that the results are very similar when we classify an NT/NH combi choice as NT, and an ES/CS choice as ES, instead of using the students' answer in the questionnaire to attribute combi profile choices to one of the four baseline study profiles. The results are also robust to treating the combi profiles as their own category, where combi profiles are ordered between the baseline study profiles, that is, NT<NT/NH<NH<ES<ES/CS<CS. Finally, the results also hold when we use each student's own ranking of his or her chosen profile (i.e. the girls rank the profiles they choose themselves lower than the boys rank their own chosen profiles).

To provide additional insights on the gender differences in profile choice, we run probit regression on choosing the most prestigious profile, NT, compared to any other profile. We compute the marginal effects evaluated at a male student and average values for the other five variables we used in Column 3 of Table 5.5. That is, the following represent gender differences when controlling for grades and feelings about mathematical prowess. We find that girls are 23 (s.e. 0.05) percentage points less likely to choose NT, a significant difference ($p = 0.00$). When we redo the exercise for choosing the least prestigious profile, CS, compared to any other profile, the marginal coefficient shows that female students are 7 (s.e. 0.03) percentage points more likely to choose CS than boys, also a significant difference ($p = 0.03$).

To summarize, boys and girls in our sample have similar grades, though boys are more confident about their mathematical prowess. Boys and girls, however, differ vastly in their study profile choices, with girls choosing significantly less prestigious profiles. This is the case even after controlling for grades and feelings of mathematical prowess. In the next section we assess the

[25] Alternatively, when we use simple OLS regressions, where NT is modeled as a choice of 1 up to CS as a choice of 4. The coefficient on female is 0.296 (s.e. 0.105, $p<0.01$) without any controls. It increases to 0.349 (s.e. 0.099, $p< 0.01$) when we add the controls from Column 2 in Table 5, which is slightly larger than the coefficient on the GPA which is -0.297 (s.e. 0.085, $p<0.01$). When we add all the controls from Column 3 the gender coefficient is 0.225 (s.e. 0.095, $p<0.05$), comparable to the coefficient on the GPA of -0.251 (s.e. 0.083, $p< 0.01$).

psychological attributes of the students in our sample. We will then assess whether these attributes can help account for gender differences in study profile choices.

5.4.2 Gender differences in competitiveness and other psychological attributes

While the students in our sample have been filtered on the basis of their score in the achievement test in primary school, all students enrolled in the pre-university track have until this stage been exposed to exactly the same curriculum. Hence, differences in competitive attitudes or other psychological attributes at this stage cannot be the result of differences in exposure to, or experiences in different study programs. More specifically, we do not have to worry that any differences between students across chosen study profiles are due to the exposure to different teachers, classmates and so on. We can therefore assess the psychological attributes of students which, in turn, may influence study profile choice.[26]

The average performance of boys in the round 1 piece rate is 6.60, significantly larger than the 5.94 of girls ($p = 0.03$ using a two-sided t-test).[27] In the round 2 tournament, there is no significant difference in performance, boys correctly solve on average 7.90 problems, compared to 7.42 for girls ($p=0.15$). Since students compete only against students in their own class, we compute for each student the chance of winning the round 2 tournament given their performance and the performance of their classmates.[28] For boys, the average chance to win the tournament is 28%, which is not significantly different from the 25% chance of girls ($p=0.23$). Provided the performance in round 3 is not lower than the performance in round 2, every student with a chance of winning the tournament of 25 percent and higher has higher expected earnings when choosing to enter the tournament in round 3. This would result in 40 percent of the boys and 36 percent of the girls entering the tournament ($p = 0.26$, Fischer's exact test). We find that 49 percent of boys and only half as many, 23 percent, of girls enter the tournament, a significant difference ($p = 0.00$, Fischer's exact test).

Table 5.6 shows marginal effects of a probit regression of tournament entry in round 3. Girls have a 26 percentage points lower probability of entering the tournament compared to boys, when controlling for performance in the round 2 tournament, the improvement in performance between

[26]We analyze the gender competition data of the 362 students in our sample. While 397 students participated in the experiment, we addressed why we had to exclude 35 of them. Those students, however, participated in the round 2 tournament of our experiment. So, while we drop them from the description of results, we, of course, use their data when calculating relative measures, such as the chance of winning the tournament, the accuracy of the guessed rank and, just as in the previous section, the relative math grade.

[27]When we use a non-parametric Mann-Whitney test, all results are basically the same.

[28]To compute the chance of winning the tournament for each participant, we use simulations and randomly draw one thousand different comparison groups of three from a participants' own class. If two performances were tied for first place, a 0.5 win was assigned (1/3 in case of three tied performances and 0.25 in case of four).

Table 5.6: Determinants of tournament entry

	(1)	(2)	(3)	(4)	(5)
Female	-0.261***	-0.191***	-0.166***	-0.153***	-0.142**
	(0.051)	(0.055)	(0.055)	(0.056)	(0.057)
Tournament	0.054***	0.022	0.020	0.024	0.014
	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)
T - PR	-0.028**	-0.022*	-0.018	-0.018	-0.014
	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)
Win Prob	0.231	0.062	0.041	-0.006	0.087
	(0.204)	(0.210)	(0.211)	(0.215)	(0.228)
Guessed rank		-0.292***	-0.293***	-0.274***	-0.269***
		(0.044)	(0.043)	(0.044)	(0.043)
Lottery			0.075***	0.033	0.037
			(0.023)	(0.024)	(0.025)
Risk-taking				0.082***	0.089***
				(0.019)	(0.019)
Math grade					0.142**
					(0.067)
GPA					-0.121*
					(0.063)
Math Relative					0.078
					(0.248)
Math quartile					0.069
					(0.044)
Math hard					-0.009
					(0.016)
N	362	362	362	362	362

Dependent variable: round 3 choice of compensation scheme (1=tournament and 0=piece rate). The table presents marginal effects of coefficients of a probit regression evaluated at a male student with a 0.25 chance of winning (the rest of the variables are evaluated at the sample mean). Standard errors of the marginal effects are in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ of the underlying coefficient.

the round 1 piece rate and the round 2 tournament, and the chance of winning the tournament. These results are very much in line with those of Niederle and Vesterlund (2007) and the large resulting literature (see Niederle and Vesterlund, 2011).

We now turn to the analysis of the two other psychological attributes, confidence and risk aversion. To elicit confidence, we asked students to rank their relative performance in the round 2 tournament from 1 (best) to 4 (worst) of their group of four. Students received € 1 if they guessed their rank correctly. The average guessed rank is 2.14 for boys, and 2.56 for girls, with the two distributions being significantly different (p=0.00; Fischer's exact test). We find that 32 percent of the boys and 11 percent of the girls believe that they are the best performers within their group, a significant difference (p = 0.00 Fischer's exact test). To assess those beliefs, we compute for each student the optimal guessed rank, that is, the guess that would have maximized their expected earnings, given the performances of the other students in their class.[29] While 56 boys and 21 girls believe they have the highest performance in their group, the numbers would be 46 and 41 if the optimal guessed rank is used. Using the optimal guessed rank, there is no gender difference in optimal beliefs, which average 2.42 for boys and 2.57 for girls (p=0.54, Fisher's exact test). Despite this, an ordered probit regression of the guessed rank on the optimal guessed rank and a female dummy delivers a female coefficient of 0.500 (s.e. 0.118, p = 0.00).[30] This confirms that girls, given their relative performance, are significantly less confident than boys.

To assess risk attitudes, we use an incentivized risk measure, which has participants choose among lotteries, where 1 is the risk free choice of €2 and 5 is the risky choice to receive €6 with a 50 percent chance.[31] Boys on average pick a significantly riskier lottery, choosing on average 3.46 compared to 2.99 for girls, a significant difference (p = 0.00; t-test). As a second risk measure we simply asked students to evaluate themselves whether they are "generally a person who is fully prepared to take risks or do you try to avoid taking risks?" The answer is on a scale from 0 ("unwilling to take risks") to 10 ("fully prepared to take risk"). Boys answer on average 6.52 compared to 5.96 for girls, a significant difference (p=0.00; t-test).[32]

The two additional psychological measures, confidence and risk aversion, may not only affect the choice of study profiles, but also of tournament entry. In Niederle and Vesterlund (2007) gender

[29] We compute the optimal guessed rank through simulation. We randomly draw a thousand different comparison groups of three from a participants' own class. We count the number of times a student ranked first, second, third and fourth. If two performances were tied for first (or second, or third) spot, 0.5 was assigned to both rank one and two (or rank two and three, or rank three and four). 1/3 was assigned in case of three tied performances and 0.25 in case of four. The mode of the ranks is the best guess as it maximizes expected earnings.

[30] The coefficient on the optimal guessed rank is 0.622 (s.e. 0.057, p=0.00).

[31] The remaining three lottery choices are 2, 3 and 4 for the 50/50 lotteries with linearly increasing riskiness and expected payoffs: 3 or 1.5; 4 or 1; 5 or 0.5.

[32] The correlation between the two risk measures is 0.42 in the whole sample, and 0.45 and 0.34 in the sub-samples of boys and girls, respectively.

differences in confidence account for about one third of the gender gap in tournament entry, while gender differences in risk attitudes only play a minor role. In their survey, Niederle and Vesterlund (2011) show that these findings are the norm in the literature. Adding the guessed rank to the probit regression on tournament entry in Table 5.6, the gender coefficient drops by about 27 percent and is now 19 percentage points and still very significant. Adding the lottery and questionnaire measures of risk to performances and beliefs on relative performance reduces the gender gap in tournament entry by a further 4 percentage points to 15 percentage points which is a reduction of 15 percent in terms of the initial gender difference (compare Columns 2 and 4). Finally, when we include measures of mathematical ability and overall grades, as well as measures of mathematical prowess, the gender gap in tournament entry is only slightly reduced, leaving a large and significant gender gap in tournament entry of 14 percentage points.

In summary, the students in our sample follow the standard gender differences in the choice of competition (see Niederle and Vesterlund, 2011). Controlling for performance, girls are about 26 percentage points less likely to enter a tournament. Boys have significantly more optimistic views about their relative performance than girls, and this gender difference in confidence accounts for about one quarter of the gender gap in tournament entry. Risk attitudes, measured by a lottery choice and a simple questionnaire item, also significantly predict tournament entry, and account for an additional 15 percent of the original gender gap in tournament entry. A final 5 percent of the original gender gap can be accounted for when controlling for grades and beliefs on mathematical prowess.

5.4.3 Can competitiveness account for gender differences in study profile choices?

The students in our sample represent a classical situation. Boys and girls agree on which academic profiles are the most prestigious and provide the highest rewards. In addition, boys and girls do not differ in math grades, and if anything, girls have slightly higher GPA's than boys. Despite these facts, girls are significantly less likely to choose the prestigious NT profile, and, in turn, significantly more likely to choose the least prestigious profile, CS. Gender differences in study profile choices remain significant when controlling for academic performance as well as mathematical prowess.

The students in our sample also exhibit the standard gender gap in competitiveness. In this section we assess whether these gender differences in competitiveness can help account for the gender gap in profile choice. This would show that competitiveness correlates with important decisions outside of the laboratory, and more importantly, that competitiveness is an important psychological attribute that can help account for gender differences in educational choices and as such labor market outcomes.

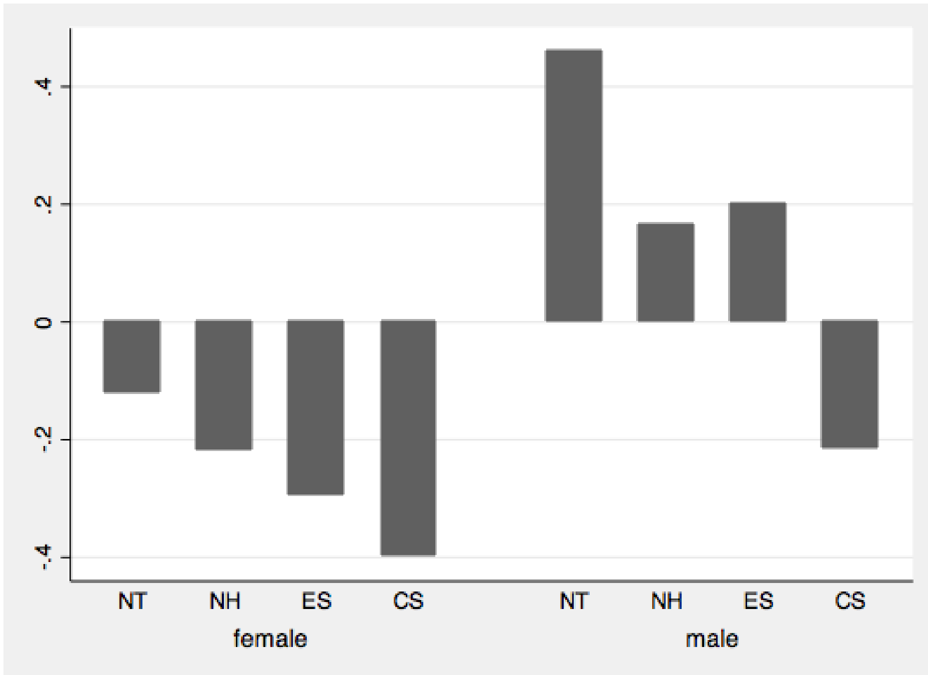
To use the decision to enter competitions as an attribute, we compute a continuous measure of competitiveness. Specifically, we run a linear regression of the decision to enter the tournament in round 3 on the round 2 tournament performance (Tournament), the increase in performance between the round 2 tournament performance and the round 1 piece rate performance (T-P) and the probability of winning the tournament in one's class (WinProb), see equation below.

$$Choice_i = \alpha_1 Tournament_i + \alpha_2 (T_i - P_i) + \alpha_3 WinProb_i + \varepsilon_i$$

We then assign each subject i the residual ε_i, normalize the resulting variable to have variance of one and call this the subject's competitiveness. For all subjects it is the case that $-1 \leq \varepsilon_i \leq 1$. Furthermore, for all 362 subjects but one is $\varepsilon_i > 0$ if and only if the subject chose to enter the tournament in round 3. Competitiveness measures how much more likely a subject is to enter a competition, given his or her performances, compared to all other subjects. Recall that tournament entry is higher the higher the performance of subjects in the addition task. Therefore, given our definition of competitiveness, a subject who enters the tournament with a high score has a relatively lower estimated competitiveness than a subject that entered the tournament with a low score. In addition, this measure is not correlated with the performance and chance of winning of the participant, and as such is a more suitable measure of competitiveness than the mere binary tournament entry variable. This is especially relevant when we regress competitiveness on study profile choice, since we do not want competitiveness to merely be another measure of (academic) performance.

In Figure 5.1, we show for each profile choice the mean competitiveness of boys and girls who chose that profile. The figure shows that for both boys and girls the students who enter the NT track are the most competitive students, followed by NH and ES, while those that enter the CS track are the least competitive ones. This ranking is even more pronounced among boys. This indicates that competitiveness as measured by our short classroom experiment may help account for the gender differences in study profile choice.

Figure 5.1: Competitiveness by gender and profile (conditional on absolute and relative performance)



We start with an intuitive way of investigating whether gender differences in competitiveness can help explain the gender difference in profile choice. We classify a student as competitive if $\varepsilon_i > 0$, which coincides with entering the tournament. We compare the impact of gender on profile choice for different subpopulations by gender and tournament entry. For example, if the gender gap in profile choice is unrelated to competitiveness, the impact of gender on profile choice should be the same for the subsample made up of competitive boys (Comp B) and non-competitive girls (N-comp G) as for the subsample made up of non-competitive boys and competitive girls.

This idea is explored in Table 5.7 which reports coefficients of regressions of profile choice on a female dummy for subsamples split by gender and competitiveness. The top part of Table 5.7 has ordered probit estimations that rank profiles NT<NH<ES<CS. The table shows that the gender gap in profile choice, which is significant for the whole sample, varies strongly with competitiveness. The gender gap in profile choice increases with the competitiveness of boys and decreases with the competitiveness of girls. When we consider competitive boys and non-competitive girls, then, on its own, gender bridges about 34% of the gap between choosing the best and the worst profile (see Column 2 of Table 5.7). When, on the other hand, we consider non-competitive boys and competitive girls, then there is no gender difference in profile choice at all. The difference in the gender effect between the groups made up of competitive boys and non-competitive girls and the other way round is significant.

Table 5.7: Gender effects by subsample

	(1) Ordered probit	(2) Female/(C3-C1)		N
(1) Comp B & n-comp. G	0.64***	0.34***		230
(2) Comp. B & comp. G	0.50**	0.28**		129
(3) N-comp. B & n-comp. G	0.14	0.08		233
(4) N-comp. B & comp. G	0.01	0.01		132
(5) Whole sample	0.32***	0.19		
P-value (1) vs (4)	0.01			

Probit (marginal effects)	(1) NT vs rest	(2) N vs S	(3) CS vs rest	(4) Best vs rest
(1) Comp B & n-comp. G	-0.33***	-0.09	0.12***	-0.32***
(2) Comp. B & comp. G	-0.30***	-0.00	0.11**	-0.24**
(3) N-comp. B & n-comp. G	-0.15***	0.05	0.02	-0.15**
(4) N-comp. B & comp. G	-0.12	0.14	0.01	-0.07
(5) Whole sample	-0.23***	0.00	0.07**	-0.22***
P-value (1) vs (4)	0.06	0.05	0.06	0.03

Coefficients are from regressions of profile choice on a female dummy only; standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; p-values are from post-estimation tests of equality of the female coefficient using Stata 11's suest command.

The probit models in the lower part of Table 5.7 give a more detailed view on this result. We first consider the probability of choosing the most prestigious profile (NT) compared to any other study profile. When we consider only competitive boys and non-competitive girls, girls are 33 percentage points less likely to choose NT. When instead we consider competitive girls and non-competitive boys, girls are only 12 percentage points less likely to choose the NT profile, and in fact, there is no significant gender difference. Furthermore, the gender difference in choosing NT is significantly smaller when we consider competitive girls and non-competitive boys than when we consider competitive boys and non-competitive girls. The results are similar when we consider choices between the Nature and the Society profiles (the top two versus the bottom two), or when we consider the probability of choosing CS, the least prestigious profile, compared to any other profile. Finally, we consider a student specific ordering of study profiles. Specifically, for each student we ask whether they pick the profile that they themselves deem to be the one chosen by the best students, or another profile. In all cases the results are very similar. Gender differences are reduced when we reduce the competitiveness of boys and increase the competitiveness of girls.

These results show that gender differences in profile choice are strongly related to gender differences in competitiveness. We now turn to detailed regressions to confirm that the impact of competitiveness on profile choice is robust when controlling for academic performance, where we consider both actual and perceived ability. Furthermore, we aim to assess what part of the gender

Table 5.8: Profile choice: ordered probit regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.342***	0.265**	0.449***	0.390***	0.204*	0.142	0.329***	0.269**
	(0.114)	(0.119)	(0.121)	(0.126)	(0.117)	(0.125)	(0.123)	(0.131)
Competitive		-0.156***		-0.141**		-0.132**		-0.146**
		(0.058)		(0.060)		(0.059)		(0.061)
Math grade			-0.188	-0.163			-0.016	0.010
			(0.133)	(0.137)			(0.141)	(0.146)
GPA			-0.360***	-0.396***			-0.326***	-0.363***
			(0.105)	(0.107)			(0.109)	(0.111)
Rel Math grade			0.509	0.501			0.345	0.341
			(0.475)	(0.487)			(0.484)	(0.500)
Math quartile					0.362***	0.366***	0.336***	0.342***
					(0.078)	(0.079)	(0.075)	(0.077)
Math difficulty					0.122***	0.119***	0.076**	0.074**
					(0.027)	(0.028)	(0.031)	(0.032)
Cut 1	-0.404	-0.450	-3.957	-4.073	0.643	0.602	-1.766	-1.879
Cut 2	0.251	0.213	-3.205	-3.315	1.435	1.400	-0.948	-1.053
Cut 3	1.367	1.341	-1.961	-2.059	2.731	2.706	0.392	0.299
Female/(C3-C1)	0.193	0.148	0.225	0.194	0.098	0.067	0.152	0.124
Dif.	23.3%		13.8%		31.6%		18.4%	
Bootstrap p-value	0.003		0.009		0.013		0.008	
Observations	362	362	362	362	362	362	362	362

Coefficients are from ordered probit regressions; bootstrap standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

difference in profile choice can be attributed to gender differences in competitiveness.

Table 5.8 shows ordered probit regressions on the ranked profile choice. Column 1 recalls the result that the being female bridges 19.3 percent of the gap between choosing the best and the worst profile. Adding competitiveness as an additional control in Column 2 significantly reduces the effect of being female by 23.3 percent down to 14.8 percent. This reduction is significant at p < 0.01.[33] The reduction in the female coefficient on profile choice is driven by competitiveness significantly pushing students into more prestigious profiles. An increase in competitiveness by one standard deviation bridges 8.7 percent of the gap between choosing the best and worst profile, and corresponds to about 59 percent of the size of the effect of being female. Over the following columns, we add controls for real and perceived ability, separately and jointly. The competitiveness measure is robust and stays significant throughout.

Pairwise comparisons between Columns 3 and 4, 5 and 6, and 7 and 8 confirm that competitiveness explains a substantial part of the gender gap in profile choice no matter which controls we include.

[33]We use bootstrap with 10'000 iterations and count the fraction of differences between Column 2's and Column 1's Female/(C3-C1) which are negative or zero.

The change in the gender coefficient upon inclusion of our competitiveness measure is significant for all specifications. When including the full set of controls in Columns 7 and 8, the reduction in the effect of being female is still 18.4 percent. This confirms that competitiveness explains a substantial part of the gender gap in profile choice.

As a robustness check, in the appendix we report ordered probit regressions for our alternative definitions of profile choice. Table 5.11 shows that the results hold when treating all NH/NT-combi students as NT students and all ES/CS-combi students as ES students. Tournament entry significantly affects the level of the chosen profile and explains a significant part of the gender gap in profile choice in all specifications. As a further robustness check, we treat the two combi profiles as separate choices. Table 5.12 shows that results again carry over. Our competitiveness measure is significant and also significantly affects the gender gap in all specifications. As a final robustness check, Table 5.13 shows that the results also hold when we use each student's own ranking of their chosen profile.

Competitiveness, as measured by the decision to enter a tournament in a classroom experiment can account for 18 percent of the gender gap in choices when controlling for grades and feelings of mathematical prowess. This shows the importance of this attribute as a determinant of gender differences in career choice. It also confirms the external validity of the large experimental literature on gender differences in tournament entry in stereotypical male tasks.

5.4.4 Decomposing the effect of competitiveness

We have seen that about 25 percent of gender differences in competitiveness, that is differences in tournament entry controlling for performance, can be attributed to gender differences in confidence. An additional 15 percent to gender differences in risk attitudes. Confidence and risk attitudes can moreover be expected to affect the choice of profile. We therefore, in an additional analysis, decompose the competitiveness measure into a measure of pure competitive attitudes, which we call *netcompete*, a measure of confidence, and a measure of risk attitudes.

Specifically, to obtain the measure of pure competitive attitudes, we run a linear regression where, in addition to controlling for performances and the chance of winning, we also control for confidence (as measured by the guessed rank) and risk preferences (as measured by the lottery choice and the risk question), see equation below.

$$Choice_i = \beta_1 Tournament_i + \beta_2(T_i - P_i) + \beta_3 WinProb_i + \beta_4 Belief_i + \beta_5 Lottery_i + \beta_6 Risk_i + \pi_i$$

We then assign each subject i the residual π_i. Then, just as in the previous section we normalize the measure such that the variance is one, and call this variable the subject's competitive attitude denoted by *netcompete*.

To have a measure of confidence which is not simply another measure of performance, we run a linear regression of the belief about the relative performance in the round 2 tournament on the round 2 tournament performance (Tournament), the increase in performance between the round 2 tournament performance and the round 1 piece rate performance (T-P), and the probability of winning the tournament in one's class (WinProb), see equation below.

$$Belief_i = \gamma_1 Tournament_i + \gamma_2 (T_i - P_i) + \gamma_3 WinProb_i + \varphi_i$$

We then assign each subject i the residual φ_i and normalize the variable to have a variance of one and call it the subject's confidence. We also multiply the residual by (-1) such that a higher confidence measure is associated with higher confidence for a given performance.

In Figure 2 we plot the average values of competitive attitudes (netcompete), confidence and the two risk attitude measures by profile choice and gender. The figure shows that netcompete is positively correlated with profile quality for both boys and girls which suggests that it is an important variable to account for gender differences in study profile choice. The same is true to a lesser extent for confidence which too might play a role in explaining the gender gap in profile choice. There is less of a clear pattern for risk preferences.

In Table 5.9, we report ordered probit regressions of profile choice on netcompete, confidence and risk attitudes. Adding these variables first separately in Columns 2 to 4 and then in various combinations in Columns 5 to 8, it becomes apparent that around half of the 25 percent reduction in the gender coefficient upon controlling for competitiveness is due to gender differences in pure competitiveness and half to gender differences in confidence. Risk attitudes play only a minor role. In Columns 9 to 13, we control for real and perceived ability. Interestingly, the effect of pure competitiveness is not at all affected while the effect of confidence drops to zero. The same is true for the effects of competitiveness and confidence on the gender gap. This indicates that the effect of competitiveness on the gender gap in profile choice conditional on real and perceived ability is due almost entirely to gender differences in pure competitiveness and not to gender differences in confidence or risk aversion.

5.5 Conclusions

This is the first study that examines whether experimentally measured gender differences in competitiveness can account for gender differences in career choices. We analyze the first important

Figure 5.2: Psychological attributes by gender and profile

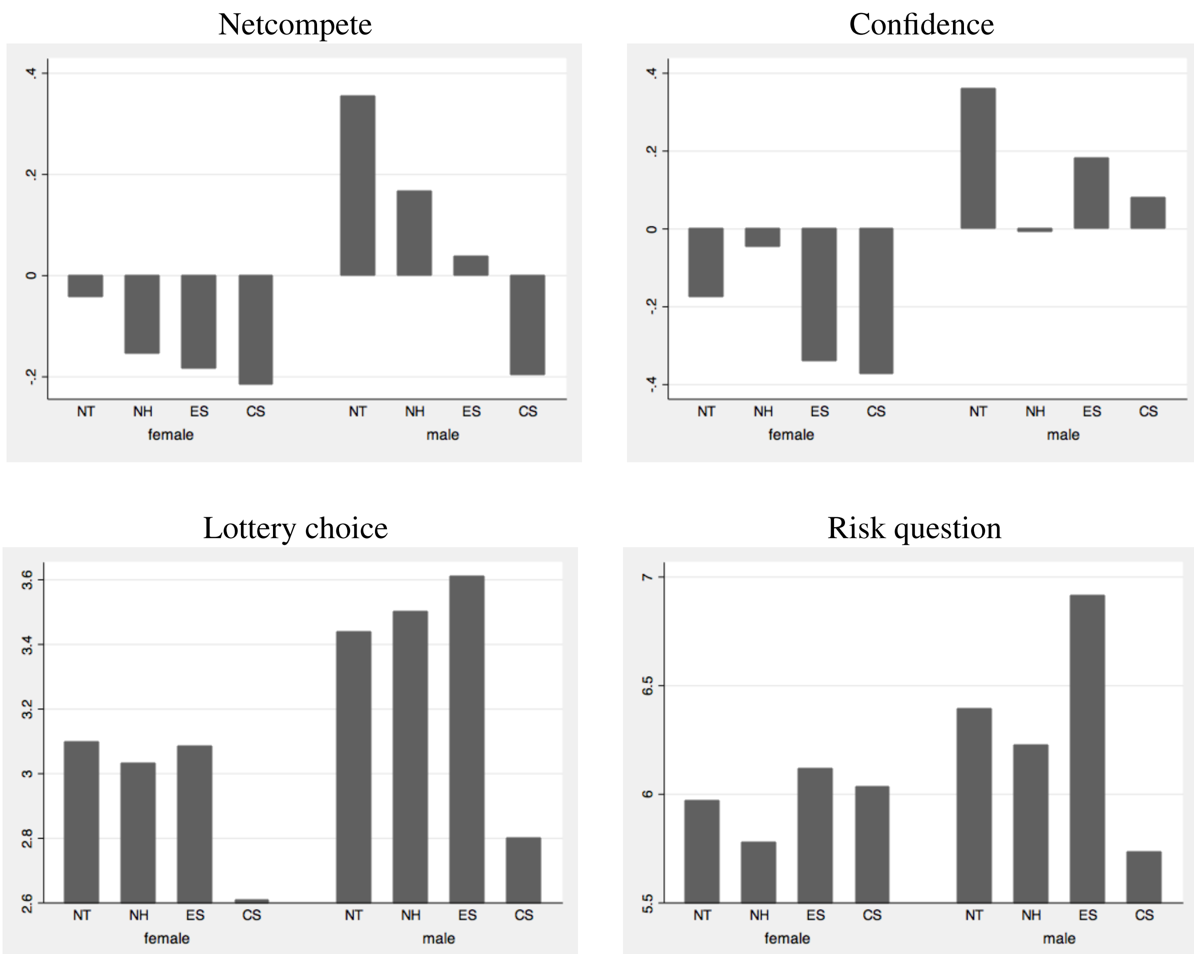


Table 5.9: Profile choice: ordered probit regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Female	0.342***	0.304***	0.298**	0.329***	0.256**	0.289**	0.286**	0.242**	0.329***	0.290**	0.342***	0.299**	0.264**
	(0.114)	(0.117)	(0.117)	(0.119)	(0.119)	(0.121)	(0.121)	(0.123)	(0.123)	(0.129)	(0.126)	(0.131)	(0.133)
Netcompete		-0.129**			-0.134**	-0.130**		-0.135**		-0.161***			-0.162***
		(0.056)			(0.057)	(0.057)		(0.057)		(0.060)			(0.061)
Confidence			0.109*		0.116*		0.120*	0.127**			-0.037		-0.022
			(0.061)		(0.061)		(0.063)	(0.064)			(0.066)		(0.069)
Lottery choice				0.040		0.039	0.051	0.051				0.038	0.034
				(0.036)		(0.036)	(0.036)	(0.037)				(0.037)	(0.038)
Risk				-0.079		-0.081	-0.082	-0.084				-0.097*	-0.099*
				(0.051)		(0.051)	(0.051)	(0.051)				(0.056)	(0.057)
Math grade									-0.016	0.017	-0.020	-0.026	0.004
									(0.141)	(0.145)	(0.147)	(0.149)	(0.148)
GPA									-0.326***	-0.363***	-0.322***	-0.310***	-0.345***
									(0.109)	(0.111)	(0.113)	(0.113)	(0.114)
Rel Math grade									0.345	0.357	0.358	0.286	0.307
									(0.484)	(0.496)	(0.498)	(0.505)	(0.502)
Math quartile									0.336***	0.352***	0.341***	0.341***	0.361***
									(0.075)	(0.077)	(0.078)	(0.078)	(0.080)
Math difficulty									0.076**	0.077**	0.077**	0.081**	0.081**
									(0.031)	(0.032)	(0.033)	(0.033)	(0.034)
Cut 1	-0.404	-0.428	-0.430	-0.417	-0.456	-0.456	-0.384	-0.422	-1.766	-1.790	-1.744	-1.814	-1.862
Cut 2	0.251	0.234	0.229	0.241	0.210	0.208	0.279	0.247	-0.948	-0.959	-0.924	-0.992	-1.027
Cut 3	1.367	1.357	1.351	1.364	1.340	1.340	1.410	1.388	0.392	0.393	0.414	0.357	0.335
f/(c3-c1)	0.193	0.170	0.167	0.185	0.143	0.161	0.159	0.134	0.152	0.133	0.158	0.138	0.120
Dif.		11.9%	13.5%	4.1%	25.9%	16.6%	17.6%	30.6%		12.5%	-3.9%	9.9%	21.1%
Bootstrap p-value	ref.	0.013	0.033	0.296	0.003	0.049	0.061	0.007	ref.	0.011	0.706	0.123	0.044
Observations	362	362	362	362	362	362	362	362	362	362	362	362	362

Coefficients are from ordered probit regressions; bootstrap standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$;

career choice that young people in the Netherlands make and for which we observe substantial gender differences. At the end of the third year, students in the pre-university track of secondary school choose between four study profiles which are ranked according to difficulty and prestige in the following order: a science-oriented profile (NT), a health-oriented profile (NH), a social science-oriented profile (ES) and a humanities-oriented profile (CS). While 40 percent of the boys in our sample choose the challenging NT profile, only 17 percent of the girls do so. Girls, on the other hand, are twice as likely to choose CS, the least demanding and prestigious of the profiles.

Ordered probit regressions confirm that girls on average pick a significantly lower ranked profile. This is despite the fact that the girls in our sample have a significantly higher GPA than the boys and do no worse at math. Prior to the moment that these students made their final choice (and were subsequently exposed to different academic subjects), we administered an experiment to elicit their competitiveness.

Like previous studies, we find that boys are more competitive than girls. This also holds when we control for differences in risk attitudes, performance, confidence and (perceived) ability. We also find that competitiveness varies strongly and significantly across chosen profiles with students picking better ranked profiles being more competitive. Ordered probit regressions confirm that competitiveness strongly affects profile choice and that this effect is robust to the inclusion of controls for grades and perceived talent for math.

Most importantly, our simple measure of competitiveness can explain around 20 percent of the gender difference in profile choice. This result is very robust, as controlling for competitiveness significantly reduces the gender gap in all our specifications.

We show that the gender gap in competitiveness is partially explained by gender differences in confidence and, to a lesser degree, risk attitudes. In a further step, we therefore decompose our measure of competitiveness into measures of pure competitiveness, confidence, and risk attitudes. We show that the effect of competitiveness on the gender gap in profile choice conditional on real and perceived ability is due almost entirely to pure competitiveness.

In summary, we show that gender gap in experimentally measured competitiveness can explain a large and significant part of the gender differences in an important academic career choice in a group of Dutch secondary school students. This demonstrates the external validity of the robust laboratory finding of gender differences in competitiveness.

5.6 Appendix

Table 5.10: Gender and profile choice (alternative specifications)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	NH/NT as NT and ES/CS and ES			NH/NT and ES/CS as separate profiles			Students' own ranking		
Female	0.275**	0.331***	0.219*	0.433***	0.494***	0.380***	0.414***	0.538***	0.460***
	(0.121)	(0.126)	(0.130)	(0.115)	(0.118)	(0.121)	(0.117)	(0.124)	(0.126)
Math Grade		-0.181	-0.031		-0.036	0.124		0.145	0.277**
		(0.145)	(0.152)		(0.127)	(0.134)		(0.124)	(0.130)
GPA		-0.383***	-0.357***		-0.319***	-0.284***		-0.367***	-0.337***
		(0.107)	(0.111)		(0.103)	(0.106)		(0.109)	(0.110)
Rel. Math Gr.		0.402	0.226		0.869*	0.692		1.225***	1.136**
		(0.494)	(0.499)		(0.462)	(0.466)		(0.466)	(0.473)
Math Quartile			0.354***			0.360***			0.150*
			(0.082)			(0.078)			(0.077)
Math Difficulty			0.058*			0.068**			0.070**
			(0.032)			(0.031)			(0.032)
Cut 1	-0.098	-3.818	-1.828	-0.568	-2.759	-0.643	-0.251	-1.330	0.246
Cut 2	0.281	-3.386	-1.358	-0.006	-2.149	0.015	0.578	-0.433	1.171
Cut 3	1.475	-2.089	0.033	0.372	-1.723	0.477	1.505	0.562	2.193
Cut 4				1.299	-0.711	1.576			
Cut 5				1.571	-0.429	1.874			
F/(Cmax-C1)	0.175	0.191	0.118	0.202	0.212	0.151	0.236	0.284	0.236
Observations	342	342	342	342	342	342	354	354	354

Coefficients are from ordered probit regressions; robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5.11: Profile choice: ordered probit regression (treating NH/NT-combi students as NT students)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.275**	0.206	0.331***	0.274**	0.115	0.065	0.219*	0.168
	(0.121)	(0.126)	(0.126)	(0.132)	(0.125)	(0.134)	(0.130)	(0.139)
Competitive		-0.142**		-0.135**		-0.113*		-0.131**
		(0.061)		(0.062)		(0.063)		(0.065)
Math grade			-0.181	-0.159			-0.031	-0.011
			(0.145)	(0.151)			(0.152)	(0.159)
GPA			-0.383***	-0.415***			-0.357***	-0.388***
			(0.107)	(0.110)			(0.111)	(0.115)
Rel Math grade			0.402	0.395			0.226	0.224
			(0.494)	(0.513)			(0.499)	(0.521)
Math quartile					0.378***	0.378***	0.354***	0.356***
					(0.084)	(0.086)	(0.082)	(0.084)
Math difficulty					0.101***	0.098***	0.058*	0.056*
					(0.028)	(0.029)	(0.032)	(0.033)
Cut 1	-0.098	-0.136	-3.818	-3.929	0.930	0.894	-1.828	-1.942
Cut 2	0.281	0.246	-3.386	-3.495	1.381	1.346	-1.358	-1.470
Cut 3	1.475	1.453	-2.089	-2.184	2.732	2.707	0.033	-0.065
f/(c3-c1)	0.175	0.130	0.191	0.157	0.064	0.036	0.118	0.090
Dif.	25.7%		17.8%		43.7%		23.7%	
Bootstrap p-value	0.010		0.013		0.038		0.021	
Observations	342	342	342	342	342	342	342	342

Coefficients are from ordered probit regressions; bootstrap standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5.12: Profile choice: ordered probit regression (treating NH/NT and ES/CS-combi as separate choices)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.433***	0.366***	0.494***	0.441***	0.289**	0.239*	0.380***	0.330**
	(0.115)	(0.119)	(0.118)	(0.124)	(0.117)	(0.125)	(0.121)	(0.129)
Competitive		-0.142**		-0.133**		-0.113**		-0.129**
		(0.057)		(0.057)		(0.057)		(0.059)
Math grade			-0.036	-0.010			0.124	0.148
			(0.127)	(0.131)			(0.134)	(0.138)
GPA			-0.319***	-0.351***			-0.284***	-0.315***
			(0.103)	(0.104)			(0.106)	(0.108)
Rel Math grade			0.869*	0.873*			0.692	0.698
			(0.462)	(0.476)			(0.466)	(0.481)
Math quartile					0.379***	0.380***	0.360***	0.362***
					(0.080)	(0.082)	(0.078)	(0.081)
Math difficulty					0.096***	0.093***	0.068**	0.066**
					(0.026)	(0.027)	(0.031)	(0.031)
Cut 1	-0.568	-0.610	-2.759	-2.836	0.365	0.326	-0.643	-0.726
Cut 2	-0.006	-0.045	-2.149	-2.224	1.014	0.976	0.015	-0.066
Cut 3	0.372	0.336	-1.723	-1.796	1.463	1.427	0.477	0.398
Cut 4	1.299	1.275	-0.711	-0.773	2.530	2.502	1.576	1.510
Cut 5	1.571	1.552	-0.429	-0.484	2.820	2.797	1.874	1.813
f/(c5-c1)	0.202	0.169	0.212	0.188	0.118	0.097	0.151	0.130
Dif.	16.3%		11.3%		17.8%		13.9%	
Bootstrap P-value	0.006		0.010		0.025		0.014	
Observations	342	342	342	342	342	342	342	342

Coefficients are from ordered probit regressions; bootstrap standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5.13: Profile choice: ordered probit regression (using students' own ranking)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.414***	0.350***	0.538***	0.485***	0.324***	0.273**	0.460***	0.406***
	(0.117)	(0.124)	(0.124)	(0.131)	(0.119)	(0.126)	(0.126)	(0.134)
Competitive		-0.127**		-0.121*		-0.104*		-0.122*
		(0.059)		(0.063)		(0.060)		(0.064)
Math grade			0.145	0.167			0.277**	0.298**
			(0.124)	(0.125)			(0.130)	(0.133)
GPA			-0.367***	-0.395***			-0.337***	-0.365***
			(0.109)	(0.110)			(0.110)	(0.113)
Rel Math grade			1.225***	1.216**			1.136**	1.127**
			(0.466)	(0.473)			(0.473)	(0.484)
Math quartile					0.174**	0.178**	0.150*	0.156**
					(0.077)	(0.079)	(0.077)	(0.079)
Math difficulty					0.094***	0.091***	0.070**	0.068**
					(0.027)	(0.028)	(0.032)	(0.032)
Cut 1	-0.251	-0.289	-1.330	-1.413	0.395	0.360	0.246	0.165
Cut 2	0.578	0.546	-0.433	-0.510	1.291	1.261	1.171	1.096
Cut 3	1.505	1.480	0.562	0.491	2.279	2.253	2.193	2.125
f/(c3-c1)	0.236	0.198	0.284	0.255	0.172	0.144	0.236	0.207
Dif.	16.1%		10.2%		16.3%		12.3%	
Bootstrap p-value	0.017		0.026		0.043		0.027	
Observations	354	354	354	354	354	354	354	354

Coefficients are from ordered probit regressions; bootstrap standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Bibliography

Albrecht, J., Björklund, A., and Vroman, S. (2003). Is there a glass ceiling in Sweden? *Journal of Labor Economics*, 21:145–177.

Altemus, M., Roca, C., Galliven, E., Romanos, C., and Deuster, P. (2001). Increased vasopressin and adrenocorticotropin responses to stress in the midluteal phase of the menstrual cycle. *The Journal of Clinical Endocrinology & Metabolism*, 86(6):2525–2530.

Altonji, J. G. and Blank, R. (1999). Race and gender in the labor market. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3C, pages 3144–3259. Elsevier Science Publishers B.V., Amsterdam.

Apicella, C. L., Dreber, A., Campbell, B., Gray, P. B., Hoffman, M., and Little, A. C. (2008). Testosterone and financial risk preferences. *Evolution and Human Behavior*, 29(6):384–390.

Apicella, C. L., Dreber, A., Gray, P. B., Hoffman, M., Little, A. C., and Campbell, B. C. (2011). Androgens and competitiveness in men. *Journal of Neuroscience, Psychology, and Economics*, 4(1):54 – 62.

Arulampalam, W., Booth, A. L., and Bryan, M. L. (2007). Is there a glass ceiling over europe? exploring the gender pay gap across the wage distribution. *Industrial and Labor Relations Review*, 60(2):163–186.

Baron-Cohen, S. (2003). *The Essential Difference*. New York: Basic Books.

Baron-Cohen, S., Lutchmaya, S., and Knickmeyer, R. (2004). *Prenatal testosterone in mind*. Cambridge: MIT Press.

Barraza, J. A. and Zak, P. J. (2009). Empathy toward strangers triggers oxytocin release and subsequent generosity. *Annals of the New York Academy of Sciences*, 1167(1):182–189.

Bartling, B., Fehr, E., Marechal, M. A., and Schunk, D. (2009). Egalitarianism and competitiveness. *American Economic Review: Papers & Proceedings*, 99(2):93–98.

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., and Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron*, 58(4):639–650.

Bellis, M. A. and Baker, R. R. (1990). Do females promote sperm competition? data for humans. *Animal Behaviour*, 40(5):997–999.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1):122–142.

Bertrand, M. (2011). New perspectives on gender. In Ashenfelter, O. and Card, D., editors, *Handbooks in Economics Handbook of Labor Economics*, volume 4, part B of *Handbook of Labor Economics*, pages 1543 – 1590. Elsevier.

Bertrand, M., Goldin, C., and Katz, L. F. (2011). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, forthcoming.

Bertrand, M. and Hallock, K. F. (2001). The gender gap in top corporate jobs. *Industrial and Labor Relations Review*, 55(1):3–21.

Bloch, M., Schmidt, P. J., Su, T.-P., Tobin, M. B., and Rubinow, D. R. (1998). Pituitary-adrenal hormones and testosterone across the menstrual cycle in women with premenstrual syndrome and controls. *Biological Psychiatry*, 43(12):897 – 903.

Bos, P. A., Terburg, D., and van Honk, J. (2010). Testosterone decreases trust in socially naive humans. *Proceedings of the National Academy of Sciences of the United States of America*, 107(22):9991–9995.

Branas-Garza, P. and Rustichini, A. (2011). Organizing effects of testosterone and economic behavior: Not just risk taking. *PLoS ONE*, 6(12):e29842.

Bröder, A. and Hohmann, N. (2003). Variations in risk taking behavior over the menstrual cycle: An improved replication. *Evolution and Human Behavior*, 24(6):391 – 398.

Brown, C. and Corcoran, M. (1997). Sex-based differences in school content and the male-female wage gap. *Journal of Labor Economics*, 15(3):pp. 431–465.

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

Burnham, T. C. (2007). High-testosterone men reject low ultimatum game offers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1623):2327–2330.

Buser, T. (2010). Handedness predicts social preferences: Evidence connecting the lab to the field. Tinbergen Institute Discussion Papers 10-119/3, Tinbergen Institute.

Buser, T. (2011). The impact of the menstrual cycle and hormonal contraceptives on competitiveness. *Journal of Economic Behavior & Organization*, forthcoming.

Buser, T. (2012). Digit ratios, the menstrual cycle and social preferences. Unpublished.

Cardenas, J. C., Dreber, A., von Essen, E., and Ranehill, E. (2011). Gender differences in competitiveness and risk taking: comparing children in colombia and sweden. *Journal of Economic Behavior and Organization*, forthcoming.

Cesarini, D., Dawes, C. T., Fowler, J. H., Johannesson, M., Lichtenstein, P., and Wallace, B. (2008). Heritability of cooperative behavior in the trust game. *Proceedings of the National Academy of Sciences of the United States of America*, 105(10):3721–3726.

Cesarini, D., Dawes, C. T., Johannesson, M., Lichtenstein, P., and Wallace, B. (2009). Genetic variation in preferences for giving and risk taking. *The Quarterly Journal of Economics*, 124(2):809–842.

Chabbert Buffet, N., Djakoure, C., Christin Maitre, S., and Bouchard, P. (1998). Regulation of the human menstrual cycle. *Frontiers in Neuroendocrinology*, 19(3):151–186.

Chavanne, T. J. and Gallup Jr, G. G. (1998). Variation in risk taking behavior among female college students as a function of the menstrual cycle. *Evolution and Human Behavior*, 19(1):27–32.

Chen, Y., Katuscak, P., and Ozdenoren, E. (2009). Why can't a woman bid more like a man? *University of Michigan, Mimeo*.

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

Coren, S. (1993). *The left-hander syndrome, the causes and consequences of left-handedness*. New York: Vintage.

Coren, S. (1995). Differences in divergent thinking as a function of handedness and sex. *American Journal of Psychology*, 108(3):311–325.

Creinin, M. D., Keverline, S., and Meyn, L. A. (2004). How regular is regular? an analysis of menstrual cycle regularity. *Contraception*, 70(4):289 – 292.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–74.

Datta Gupta, N., Poulsen, A., and Villeval, M.-C. (2011). Gender matching and competitiveness: experimental evidence. *Economic Inquiry*, forthcoming.

de Wit, H., Schmitt, L., Purdy, R., and Hauger, R. (2001). Effects of acute progesterone administration in healthy postmenopausal women and normally-cycling women. *Psychoneuroendocrinology*, 26(7):697–710.

Denny, K. and O'Sullivan, V. (2007). The Economic Consequences of Being Left-Handed: Some Sinister Results. *Journal of Human Resources*, XLII(2):353–374.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants and behavioral consequences. *Journal of the European Economic Association*, 9:522–550.

Dreber, A. and Hoffman, M. (2007). Risk preferences are partly predetermined. Unpublished.

Driessen, G. and Van Langen, A. (2010). De onderwijsachterstand van jongens. Technical report, Radboud Universiteit. Nijmegen, The Netherlands.

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

Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M., and Fehr, E. (2010). Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature*, 463(7279):356–359.

Ellison, G. and Swanson, A. (2010). The gender gap in secondary school mathematics at high achievement levels: Evidence from the american mathematics competitions. *The Journal of Economic Perspectives*, 24(2):109–128.

Ermisch, A., Barth, T., Ruhle, H., Skopkova, J., Hrbas, P., and Landgraf, R. (1985). On the blood-brain barrier to peptides: accumulation of labelled vasopressin, desglynh2-vasopressin and oxytocin by brain regions. *Endocrinological Experiments*, 19(1):29 – 37.

Faurie, C., Vianey-Liaud, N., and Raymond, M. (2006). Do left-handed children have advantages regarding school performance and leadership skills? *Laterality*, 11(1):57–70.

Fehr, E. (2009). On the economics and biology of trust. *Journal of the European Economic Association*, 7(2-3):235–266.

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

Gächter, S., Herrmann, B., and Töhni, C. (2010). Culture and cooperation. *Philos Trans R Soc Lond B Biol Sci*, 365(1553):2651–2661.

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

Geschwind, N. and Galaburda, A. S. (1987). *Cerebral lateralisation*. Cambridge: MIT Press.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L. (2000). Measuring trust. *The Quarterly Journal of Economics*, 115(3):811–846.

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

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, 118(3):1049–1074.

Goldberg, E., Harner, R., Lovell, M., Podell, K., and Riggio, S. (1994). Cognitive bias, functional cortical geometry, and the frontal lobes: Laterality, sex, and handedness. *Journal of Cognitive Neuroscience*, 6(3):276–296.

Goldin, C., Katz, L. F., and Kuziemko, I. (2006). The homecoming of american college women: The reversal of the college gender gap. *The Journal of Economic Perspectives*, 20(4):133–156.

Grimshaw, G., Bryden, M. P., and Finegan, J.-A. (1995). Relations between prenatal testosterone and cerebral lateralization in children. *Neuropsychology*, 9(1):68 – 79.

Grogger, J. and Eide, E. (1995). Changes in college skills and the rise in the college wage premium. *The Journal of Human Resources*, 30(2):280–310.

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organizationn*, 3(4):367–388.

Hampson, E. and Kimura, D. (1992). Sex differences and hormonal influences on cognitive function in humans. In Becker, J. B., Breedlove, S. M., and Crews, D., editors, *Behavioral Endocrinology*, pages 357–398. MIT Press.

Hampson, E. and Young, E. A. (2008). Methodological issues in the study of hormone-behavior relations in humans: Understanding and monitoring the menstrual cycle. In Becker, J. B., Berkley, K. J., Geary, N., Hampson, E., Herman, J. P., and Young, E., editors, *Sex differences in the brain. From genes to behavior.*, pages 63–78. Oxford University Press.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., and McElreath, R. (2001). In search of homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review*, 91(2):73–78.

Herrmann, B., Töhni, C., and Gächter, S. (2008). Antisocial Punishment Across Societies. *Science*, 319(5868):1362–1367.

Hines, M. (2009). Gonadal hormones and sexual differentiation of human brain and behavior. In Pfaff, I. D. W., Arnold, A. P., Etgen, A. M., Fahrbach, S. E., and Rubin, R. T., editors, *Hormones, Brain and Behavior*, pages 1869–1909. San Diego, CA: Academic Press.

Hines, M. (2011). Gender development and the human brain. *Annual Review of Neuroscience*, 34(1):69–88.

Hines, M., Chiu, L., McAdams, L. A., Bentler, P., and Lipcamon, J. (1992). Cognition and the corpus callosum: Verbal fluency, visuospatial ability, and language lateralization related to midsagittal surface areas of callosal subregions. *Behavioral Neuroscience*, 106(1):3 – 14.

Hoffman, M. and Gneezy, U. (2010). Left handed women are more competitive than right handed men: On the biological basis of gender differences in competitiveness. *UC San Diego, Mimeo*.

Holm, H. J. and Danielson, A. (2005). Tropic trust versus nordic trust: Experimental evidence from tanzania and sweden. *Economic Journal*, 115(503):505–532.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655.

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

Israel, S., Lerer, E., Shalev, I., Uzefovsky, F., Riebold, M., Laiba, E., Bachner-Melman, R., Maril, A., Bornstein, G., Knafo, A., and Ebstein, R. P. (2009). The oxytocin receptor (oxtr) contributes to prosocial fund allocations in the dictator game and the social value orientations task. *PLoS ONE*, 4:e5535.

Johnston, D. W., Shah, M., and Shields, M. A. (2007). Handedness, Time Use and Early Childhood Development. Discussion Paper 4300, Institute for the Study of Labor (IZA).

Jones, B., Little, A., Boothroyd, L., DeBruine, L., Feinberg, D., Law Smith, M., Cornwell, R., Moore, F., and Perrett, D. (2005). Commitment to relationships and preferences for femininity and apparent health in faces are strongest on days of the menstrual cycle when progesterone level is high. *Hormones and Behavior*, 48(3):283–290.

Kastlunger, B., Dressler, S. G., Kirchler, E., Mittone, L., and Voracek, M. (2010). Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2d:4d). *Journal of Economic Psycholy*, 31(4):542–552.

Knafo, A., Israel, S., Darvasi, A., Bachner-Melman, R., Uzefovsky, F., Cohen, L., Feldman, E., Lerer, E., Laiba, E., Raz, Y., Nemanov, L., Gritsenko, I., Dina, C., Agam, G., Dean, B., Bornstein, G., and Ebstein, R. P. (2008). Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor rs3 promoter region and correlation between rs3 length and hippocampal mrna. *Genes, Brain and Behavior*, 7(3):266–275.

Knecht, S., Deppe, M., Drager, B., Bobe, L., Lohmann, H., Ringelstein, E.-B., and Henningsen, H. (2000). Language lateralization in healthy right-handers. *Brain*, 123(1):74–81.

Knickmeyer, R., Baron-Cohen, S., Raggatt, P., and Taylor, K. (2005). Foetal testosterone, social relationships, and restricted interests in children. *The Journal of Child Psychology and Psychiatry*, 46(2):198 – 210.

Kol, S. and Homburg, R. (2008). Change, change, change: hormonal actions depend on changes in blood levels. *Human Reproduction*, 23(5):1004–1006.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., and Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435(2):673–676.

Ledyard, J. (1995). Public goods: A survey of experimental research. In Kagel, J. H. and Roth, A. E., editors, *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.

LeFevre, J.-A., Kulak, A. G., and Heymans, S. L. (1992). Factors influencing the selection of university majors varying in mathematical content. *Canadian Journal of Behavioural Science*, 24(3):276–289.

Levitt, S. D. and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2):153–174.

Lonnqvist, J.-E., Verkasalo, M., Walkowitz, G., and Wichardt, P. C. (2010). Measuring individual risk attitudes in the lab: Task or ask? an empirical comparison. Unpublished.

Lust, J., Geuze, R., de Beek, C. V., Cohen-Kettenis, P., Groothuis, A., and Bouma, A. (2010). Sex specific effect of prenatal testosterone on language lateralization in children. *Neuropsychologia*, 48(2):536 – 540.

Lutchmaya, S. and Baron-Cohen, S. (2002). Human sex differences in social and non-social looking preferences, at 12 months of age. *Infant Behavior and Development*, 25(3):319 – 325.

Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., and Manning, J. T. (2004). 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development*, 77(1-2):23 – 28.

Manning, J. T. (2002). *Digit ratio. A pioneer to fertility, behavior and health*. Rutgers University Press.

Manning, J. T., Barley, L., Walton, J., Lewis-Jones, D., Trivers, R., Singh, D., Thornhill, R., Rohde, P., Bereczkei, T., Henzi, P., Soler, M., and Szwed, A. (2000). The 2nd:4th digit ratio, sexual dimorphism, population differences, and reproductive success: evidence for sexually antagonistic genes? *Evolution and Human Behavior*, 21(3):163 – 183.

Manning, J. T., Baron-Cohen, S., Wheelwright, S., and Sanders, S. (2001). The 2nd to 4th digit ratio and autism. *Developmental Medicine & Child Neurology*, 43(3):160–164.

Manning, J. T., Bundred, P. E., Newton, D. J., and Flanagan, B. F. (2003). The second to fourth digit ratio and variation in the androgen receptor gene. *Evolution and Human Behavior*, 24(6):399 – 405.

Manning, J. T. and Peters, M. (2009). Digit ratio (2d:4d) and hand preference for writing in the bbc internet study. *Laterality*, 14(5):528–540.

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

Manning, J. T. and Taylor, R. P. (2001). Second to fourth digit ratio and male ability in sport: implications for sexual selection in humans. *Evolution and Human Behavior*, 22(1):61 – 69.

Mansour, D. (2006). Use of the new progestogens in contraception and gynaecology. *The Obstetrician & Gynaecologist*, (8):229–234.

McManus, C. (2002). *Right hand, left hand. The origins of asymmetry in brains, bodies, atoms and cultures*. London: Weidenfeld & Nicolson.

Miller, G., Tybur, J. M., and Jordan, B. D. (2007). Ovulatory cycle effects on tip earnings by lap dancers: economic evidence for human estrus? *Evolution and Human Behavior*, 28(6):375 – 381.

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

Millet, K. and Dewitte, S. (2006). Second to fourth digit ratio and cooperative behavior. *Biological Psychology*, 71(1):111 – 115.

Millet, K. and Dewitte, S. (2009). The presence of aggression cues inverts the relation between digit ratio (2d:4d) and prosocial behaviour in a dictator game. *British Journal of Psychology*, 100:151 – 162.

Mobius, M., Niederle, M., Niehaus, P., and Rosenblat, T. (2011). Managing self-confidence: Theory and experimental evidence. Unpublished.

Morhenn, V. B., Park, J. W., Piper, E., and Zak, P. J. (2008). Monetary sacrifice among strangers is mediated by endogenous oxytocin release after physical contact. *Evolution and Human Behavior*, 29(6):375 – 383.

Murnane, R. J., Willett, J. B., Duhaldeborde, Y., and Tyler, J. H. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management*, 19(4):547–568.

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

Niederle, M. and Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, 3(1):601–630.

Oosterbeek, H., Sloof, R., and van de Kuilen, G. (2004). Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics*, 7(2):171–188.

Paglin, M. and Rufolo, A. M. (1990). Heterogeneous human capital, occupational choice, and male-female earnings differences. *Journal of Labor Economics*, 8(1):123–144.

Pearson, M. and Schipper, B. C. (2011a). Menstrual cycle and competitive bidding. Unpublished.

Pearson, M. and Schipper, B. C. (2011b). The visible hand: Finger ratio (2d:4d) and competitive bidding. *Experimental Economics*, forthcoming.

Penton-Voak, I. S. and Perrett, D. I. (2000). Female preference for male faces changes cyclically: Further evidence. *Evolution and Human Behavior*, 21(1):39–48.

Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burr, D. M., Murray, L. K., and Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature*, (6738):741–742.

Powell, M. and Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology*, 18(6):605–628.

Putz, D. A., Gaulin, S. J., Sporter, R. J., and McBurney, D. H. (2004). Sex hormones and finger length: What does 2d:4d indicate? *Evolution and Human Behavior*, 25(3):182 – 199.

Richardson, J. T. (1992). The menstrual cycle, cognition, and paramenstrual symptomatology. In Richardson, J. T., editor, *Cognition and the Menstrual Cycle*. Springer-Verlag.

Rivera, R., Yacobson, I., and Grimes, D. (1999). The mechanism of action of hormonal contraceptives and intrauterine contraceptive devices. *American Journal of Obstetrics & Gynecology*, 181:1263–1269.

Robinson, S. J. and Manning, J. T. (2000). The ratio of 2nd to 4th digit length and male homosexuality. *Evolution and Human Behavior*, 21(5):333–345.

Roth, A. E. (1995). Bargaining experiments. In Kagel, J. H. and Roth, A. E., editors, *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.

Ruebeck, C. S., Harrington, J. E., and Moffitt, R. (2007). Handedness and earnings. *Laterality*, 12(2):101–120.

Salonia, A., E. Nappib, R., Pontilloc, M., Daverioc, R., Smeraldid, A., Brigantia, A., Fabbria, F., Zannia, G., Rigattia, P., and Montorsia, F. (2005). Menstrual cycle-related changes in plasma oxytocin are relevant to normal sexual function in healthy women. *Hormones and Behavior*, 47:164–169.

Sanchez-Pages, S. and Turiegano, E. (2010). Testosterone, facial symmetry and cooperation in the prisoners' dilemma. *Physiology & Behavior*, 99(3):355 – 361.

Sanders, B., Wilson, J. R., and Vandenberg, S. G. (1982). Handedness and spatial ability. *Cortex*, 18(1):79–90.

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

Schipper, B. C. (2011a). Sex hormones and choice under risk. Unpublished.

Schipper, B. C. (2011b). Sex hormones and competitive bidding. Unpublished.

Small, C. M., Manatunga, A. K., and Marcus, M. (2007). Validity of self-reported menstrual cycle length. *Annals of Epidemiology*, 17(3):163 – 170.

Stern, K. and McClintock, M. K. (1998). Regulation of ovulation by human pheromones. *Nature*, 1392(6672):177 – 179.

Sutter, M. and Rützler, D. (2010). Gender differences in competition emerge early in life. Discussion Paper 5015, Institute for the Study of Labor (IZA).

Thornhill, R. and Gangestad, S. W. (1999). The scent of symmetry: A human sex pheromone that signals fitness? *Evolution and Human Behavior*, 20(3):175–201.

van den Bergh, B. and Dewitte, S. (2006). Digit ratio (2D:4D) moderates the impact of sexual cues on men's decisions in ultimatum games. *Proceedings of the Royal Society B: Biological Sciences*, 273(1597):2091–2095.

Wallace, B., Cesarini, D., Lichtenstein, P., and Johannesson, M. (2007). Heritability of ultimatum game responder behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 104(40):15631–15634.

Weinberger, C. J. (1999). Mathematical college majors and the gender gap in wages. *Industrial Relations: A Journal of Economy and Society*, 38(3):407–413.

Weinberger, C. J. (2001). Is teaching more girls more math the key to higher wages? In King, M. C., editor, *Squaring Up: Policy Strategies to Raise Women's Incomes in the United States*, chapter 11. University of Michigan Press.

Weinberger, C. J. (2005). Is the Science and Engineering Workforce Drawn from the Far Upper Tail of the Math Ability Distribution? Unpublished.

Witelson, S. F. (1985). The brain connection: The corpus callosum is larger in left-handers. *Science*, 229(4714):665–668.

Witelson, S. F. (1991). Neural sexual mosaicism: Sexual differentiation of the human temporoparietal region for functional asymmetry. *Psychoneuroendocrinology*, 16(1-3):131 – 153.

Witelson, S. F. and Goldsmith, C. H. (1991). The relationship of hand preference to anatomy of the corpus callosum in men. *Brain Research*, 545(1-2):175 – 182.

Witelson, S. F. and Nowakowski, R. S. (1991). Left out axoms make men right: A hypothesis for the origin of handedness and functional asymmetry. *Neuropsychologia*, 29(4):327 – 333.

Wozniak, D., Harbaugh, W. T., and Mayr, U. (2010). Choices About Competition: Differences by gender and hormonal fluctuations, and the role of relative performance feedback. Unpublished.

Zak, P. J. (2011). The physiology of moral sentiments. *Journal of Economic Behavior & Organization*, 77(1):53 – 65.

Zak, P. J., Kurzban, R., Ahmadi, S., Swerdloff, R. S., Park, J., Efremidze, L., Redwine, K., Morgan, K., and Matzner, W. (2009). Testosterone administration decreases generosity in the ultimatum game. *PLoS One*, 4(12).

Zak, P. J., Stanton, A. A., and Ahmadi, S. (2007). Oxytocin increases generosity in humans. *PLoS One*, 2(11).

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

Samenvatting (Summary in Dutch)

Dit proefschrift gaat over individuele verschillen in economische voorkeuren. Meestal worden voorkeuren in de economische literatuur als gegeven genomen. Economen gebruiken ze als de bouwstenen van hun modellen, schatten ze in laboratoriumexperimenten, en correleren ze met individuele uitkomsten. Maar er wordt maar zelden naar hun oorsprong gevraagd. Hoe komt het dat voorkeuren zoals risicoaversie, reciprociteit en altruïsme zo sterk tussen individuen verschillen? Zijn ze biologisch bepaald of kunnen ze worden beïnvloed door opvoeding en onderwijs?

Deze vraag is belangrijk. Zoals hoofdstuk 5 van dit proefschrift laat zien, beïnvloeden economische voorkeuren belangrijke beslissingen in ons leven, met inbegrip van de keuzes van studie en carrière. Daarom is het belangrijk om te weten hoe deze voorkeuren tot stand komen. Hoofdstukken 2 tot 4 gaan daarover. In deze hoofdstukken kijk ik naar de associatie tussen fysieke markers van hormonale en neurologische variabelen en keuzes in laboratoriumexperimenten. Een significante correlatie kan een indicatie zijn dat biologische factoren een rol spelen bij het bepalen van economische voorkeuren.

In hoofdstuk 2 onderzoek ik de invloed van hormonale fluctuaties op de competitiviteit van vrouwen. Eerdere experimenten hebben laten zien dat een groot deel van de mannelijke proefpersonen wordt aangetrokken tot competitie, terwijl een groot deel van vrouwelijke proefpersonen competitie juist uit de weg gaat. Het is voor de hand liggend om dit resultaat te extrapoleren naar keuzes op de arbeidsmarkt: het verschil in competitiviteit tussen mannen en vrouwen zou een verklaring kunnen zijn voor verschillen in carrièrekeuzes. Maar waarom worden mannen meer aangetrokken tot competitieve situaties dan vrouwen? Komt het door verschillen in opvoeding en onderwijs, of is er een aangeboren verschil? Een belangrijk biologisch verschil tussen mannen en vrouwen is de blootstelling aan geslachtshormonen. Testosteronconcentraties zijn veel hoger bij mannen, terwijl vrouwen aan veel hogere hoeveelheden van oestrogeen en progesteron worden blootgesteld. Kunnen deze hormonale verschillen een deel van het verschil in competitiviteit tussen mannen en vrouwen verklaren?

Om deze vraag te beantwoorden, maak ik gebruik van de fluctuaties in hormoonconcentraties die vrouwen gedurende de menstruatiecyclus ervaren. Vrouwen die geen hormonale anticonceptiemiddelen gebruiken, hebben te maken met sterke en voorspelbare fluctuaties in progesteron en oestrogeen (plus een aantal andere hormonen). Ook vrouwen die wel hormonale anticonceptiemiddelen gebruiken, ervaren dergelijke fluctuaties. De concentratie van oestrogeen en progesteron daalt namelijk sterk tijdens de zevendaagse onderbreking van het gebruik van de pil. Met behulp van een eenvoudig experimenteel design, waarbij proefpersonen moeten kiezen of ze voor het maken van optelsommen willen worden beloond volgens een stukloon of een toernooi, laat ik zien dat de hormoonspiegel een goede voorspeller is van de kans dat een bepaald beloningssysteem wordt gekozen. Pilgebruiksters zijn ongeveer twee keer zo competitief tijdens de zevendaagse onderbreking (wanneer hun hormoonspiegels laag zijn) dan gedurende de 21 dagen waarin ze de pil wel innemen (wanneer hun hormoonspiegels hoog zijn). Vrouwen met een natuurlijke cyclus zijn ongeveer de helft zo competitief in de dagen direct na de ovulatie dan tijdens de rest van de cyclus. Samen wijzen deze uitkomsten op een negatief effect van het hormoon progesteron op competitiviteit.

Evenals competitiviteit verschillen ook sociale voorkeuren zoals vertrouwen, altruïsme en reciprociteit sterk tussen individuen. Ook hier kan men zich afvragen in hoeverre deze verschillen door biologische factoren worden bepaald. Hoofdstukken 3 en 4 rapporteren onderzoek waarin ik proefpersonen aan een reeks economische spellen laat deelnemen om hun sociale voorkeuren te meten. Deze spellen zijn het “trust game”, het “ultimatum game”, het “public good game” en het “dictator game”. Ik correleer de keuzes die proefpersonen in deze spellen maken met lichamelijke markers van onderliggende neurologische en hormonale variabelen. Een significante correlatie wijst op een gedeeltelijke biologische basis voor sociale voorkeuren.

Hoofdstuk 3 gaat over de effecten van blootstelling aan hormonen die voor de geboorte in de baarmoeder plaatsvindt. De mate waarin we in de baarmoeder aan de hormonen testosteron en oestrogeen worden blootgesteld, varieert sterk tussen individuen, zowel voor mannen als voor vrouwen. Er wordt vermoed dat de mate van blootstelling blijvende effecten heeft op de structuur van de hersenen en daarmee op gedrag en voorkeuren op latere leeftijd. Mensen die in de baarmoeder aan een hoge concentratie van testosteron zijn blootgesteld, kunnen zich bijvoorbeeld slecht in andere mensen inleven maar zijn goed in het omgaan met systemen. De mate van blootstelling aan testosteron en oestrogeen in de baarmoeder bepaalt ook de verhouding van de lengten van de wijsvinger en ringvinger. Die verhouding is vanaf de geboorte stabiel. Door proefpersonen te vragen of hun wijsvinger langer is dan hun ringvinger, krijg ik een ruwe maatstaf voor de neurologische verschillen die gerelateerd zijn met de sterkte van de prenatale hormonale blootstelling. Ik vind dat personen met een verhouding van vingerlengten die op een sterkere blootstelling aan testosteron wijst, bij alle spellen minder genereus zijn dan anderen. Dit kan komen door een negatief effect van de prenatale blootstelling aan testosteron op empathie. Verder stel ik de vrouwelijke proefpersonen een aantal vragen over hun menstruatiecyclus en vind dat sociale voorkeuren, net als de competitiviteit, over de menstruatiecyclus variëren.

In hoofdstuk 4 correleer ik dezelfde gegevens met linkshandigheid. Linkshandige mensen verschillen in neurologisch opzicht van rechtshandige mensen (en de verschillen variëren tussen mannen en vrouwen). Ik vind dat linkshandige mannen aanzienlijk genereuzer zijn in spellen waar ontvangers de mogelijkheid hebben om zich erkentelijk te tonen, en dat zij zelf sterkere positieve reciprociteit tonen dan rechtshandige mannen. Linkshandige vrouwen zijn daarentegen aanzienlijk minder altruïstisch dan rechtshandige vrouwen. Informatie over linkshandigheid is ook beschikbaar in enquêtegegevens en ik kan dus nagaan of deze resultaten geëxtrapoleerd kunnen worden naar gedrag buiten het lab. Gebruikmakend van Nederlandse en Amerikaanse gegevens vind ik significante effecten van handvoorkeur op altruïstisch gedrag, vertrouwen in anderen, het doneren van geld en het doen van vrijwilligerswerk. In combinatie met de resultaten van hoofdstuk 3 wijst dit op een neurale basis van sociale voorkeuren.

Hoofdstuk 5 gaat niet over de oorsprong maar over de gevolgen van individuele verschillen in economische voorkeuren. We gaan dieper in op het verschil in competitiviteit tussen mannen en vrouwen. De studie die dit verschil als eerste aantoonde, heeft tot veel vervolgstudies geleid. De meeste van deze studies vinden plaats in het laboratorium. De resultaten ervan worden echter geëxtrapoleerd naar de arbeidsmarkt. De redenering is dan dat als vrouwen competitie uit de weg gaan, ze daardoor waarschijnlijk andere carrières kiezen dan mannen. Dit zou deels de verschillen tussen mannen en vrouwen op de arbeidsmarkt kunnen verklaren. Wij onderzoeken de geldigheid van deze extrapolatie door een experimentele maat voor competitiviteit te correleren met de eerste belangrijke beroepskeuze van vwo-scholieren in Nederland. In de derde klas moeten deze leerlingen een keuze maken uit vier studieprofielen: natuur & techniek, natuur & gezondheid, economie & maatschappij en cultuur & maatschappij. De profielkeuze is sterk gecorreleerd met de waarschijnlijkheid om later naar de universiteit te gaan, en met het gekozen universitaire traject. De keuzes van jongens en meisjes vertonen duidelijke verschillen. Jongens kiezen vaak voor natuur & techniek. Dit profiel wordt algemeen als het meest prestigieuze en uitdagende profiel wordt beschouwd. Meisjes kiezen vaker de minder prestigieuze natuur & gezondheid en cultuur & maatschappij profielen. Het resultaat dat jongens veel competitiever zijn dan meisjes wordt ook in onze studie bevestigd. We vinden verder dat de competitiviteit van de scholieren sterk met hun profielkeuzes correleert. Het verschil in competitiviteit tussen jongens en meisjes verklaart ongeveer 20 procent van hun verschil in studiekeuzes. Het suggereert dat verschillen in competitiviteit inderdaad gedeeltelijk verantwoordelijk zijn voor de verschillende carrièrepaden van mannen en vrouwen.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

488 B. SCHWAAB, Credit Risk and State Space Methods
489 N. BASTÜRK, Essays on parameter heterogeneity and model uncertainty
490 E.GUTIÉRREZ PUIGARNAU, Labour markets, commuting and company cars
491 M.W. VORAGE, The Politics of Entry
492 A.N. HALSEMA, Essays on Resource Management: Ownership, Market Structures and Exhaustibility
493 R.E. VLAHU, Three Essays on Banking
494 N.E. VIKANDER, Essays on Teams and the Social Side of Consumption
495 E. DEMIREL, Economic Models for Inland Navigation in the Context of Climate Change
496 V.A.C. VAN DEN BERG, Congestion pricing with Heterogeneous travellers
497 E.R. DE WIT, Liquidity and Price Discovery in Real Estate Assets
498 C. LEE, Psychological Aspects of the Disposition Effect: An Experimental Investigation
499 MHA. RIDHWAN, Regional Dimensions of Monetary Policy in Indonesia
500 J. GARCÍA, The moral herd: Groups and the Evolution of Altruism and Cooperation
501 F.H. LAMP, Essays in Corporate Finance and Accounting
502 J. SOL, Incentives and Social Relations in the Workplace
503 A.I.W. HINDRAYANTO, Periodic Seasonal Time Series Models with applications to U.S. macroeconomic data
504 J.J. DE HOOP, Keeping Kids in School: Cash Transfers and Selective Education in Malawi
505 O. SOKOLINSKIY, Essays on Financial Risk: Forecasts and Investor Perceptions
506 T. KISELEVA, Structural Analysis of Complex Ecological Economic Optimal Management Problems
507 U. KILINC, Essays on Firm Dynamics, Competition and Productivity
508 M.J.L. DE HEIDE, R&D, Innovation and the Policy Mix
509 F. DE VOR, The Impact and Performance of Industrial Sites: Evidence from the Netherlands
510 J.A. NON, Do ut Des: Incentives, Reciprocity, and Organizational Performance
511 S.J.J. KONIJN, Empirical Studies on Credit Risk
512 H. VRIJBURG, Enhanced Cooperation in Corporate Taxation
513 P.ZEPPINI, Behavioural Models of Technological Change
514 P.H.STEFFENS, It's Communication, Stupid! Essays on Communication, Reputation and

(Committee) Decision-Making
515 K.C. YU, Essays on Executive Compensation - Managerial Incentives and Disincentives
516 P. EXTERKATE, Of Needles and Haystacks: Novel Techniques for Data-Rich Economic Forecasting
517 M. TYSZLER, Political Economics in the Laboratory
518 Z. WOLF, Aggregate Productivity Growth under the Microscope
519 M.K. KIRCHNER, Fiscal Policy and the Business Cycle – The Impact of Government Expenditures, Public Debt, and Sovereign Risk on Macroeconomic Fluctuations
520 P.R. KOSTER, The cost of travel time variability for air and car travelers
521 Y.ZU, Essays of nonparametric econometrics of stochastic volatility
522 B.KAYNAR, Rare Event Simulation Techniques for Stochastic Design Problems in Markovian Setting
523 P. JANUS, Developments in Measuring and Modeling Financial Volatility
524 F.P.W. SCHILDER, Essays on the Economics of Housing Subsidies
525 S. M MOGHAYER, Bifurcations of Indifference Points in Discrete Time Optimal Control Problems
526 C. ÇAKMAKLI, Exploiting Common Features in Macroeconomic and Financial Data
527 J. LINDE, Experimenting with new combinations of old ideas
528 D. MASSARO, Bounded rationality and heterogeneous expectations in macroeconomics
529 J. GILLET, Groups in Economics
530 R. LEGERSTEE, Evaluating Econometric Models and Expert Intuition
531 M.R.C. BERSEM, Essays on the Political Economy of Finance
532 T. WILLEMS, Essays on Optimal Experimentation
533 Z. GAO, Essays on Empirical Likelihood in Economics
534 J. SWART, Natural Resources and the Environment: Implications for Economic Development and International Relations
535 A. KOTHIYAL, Subjective Probability and Ambiguity
536 B. VOOGT, Essays on Consumer Search and Dynamic Committees
537 T. DE HAAN, Strategic Communication: Theory and Experiment