

The Menstrual Cycle and Performance Feedback Alter Gender  
Differences in Competitive Choices<sup>1</sup>

*Forthcoming in The Journal of Labor Economics*

David Wozniak<sup>2</sup>

William T. Harbaugh<sup>3</sup>

Ulrich Mayr<sup>4</sup>

December 9, 2012

<sup>1</sup>The authors acknowledge support from NIA grant R21-AGR21AG030783 and NSF grant SES-1063561. We are grateful for support and direction from Chris Minson and Paul Kaplan of the University of Oregon Human Physiology Department. Seminar participants at the (June) 2009 ESA and (July) 2009 IAREP/SABE meetings provided useful ideas and comments. Correspondence should be addressed to David Wozniak (Eastern Michigan University) at [dwozniak@gmail.com](mailto:dwozniak@gmail.com). We would also like to thank Jack Wozniak for his programming assistance and Cathleen Leue, the director of the Social Science Instructional Labs and her staff at the University of Oregon. A final thanks to Sarah Wunderlich for all her help scheduling and running experiments.

<sup>2</sup>Eastern Michigan University

<sup>3</sup>University of Oregon

<sup>4</sup>University of Oregon

# The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices.

December 9, 2012

*Forthcoming in The Journal of Labor Economics*

## **Abstract**

We use a within-subjects experiment with math and word tasks to show that relative performance feedback moves high ability females towards more competitive forms of compensation and moves low ability men towards less competitive forms and eliminates gender differences in choices. We also examine females across the menstrual cycle, and find that women in the high-hormone phase are more willing to compete than women in the low phase. There are no significant differences between choices after subjects receive feedback. Thus, biological differences lead to economically significant differences, but the impact of those differences can be lowered through relative performance information.

## Introduction

Economic experiments with students have shown that when given the choice between piece rate and winner-take-all tournament style compensation, women are more reluctant than men to choose tournaments (Niederle and Vesterlund 2007). This difference has been replicated with a large scale field experiment that includes employed adults (Mayr et al. 2012). The gender difference for competitions has been used as a potential explanation for the lack of women at the top ranks in many competitive careers.

These studies all relied on a framework where subjects were not informed of their abilities relative to potential competitors. We use a within-subjects design and replicate these previous findings for a math task, and show they also exist for a word task. We then show that feedback about relative performance moves high ability females towards more competitive compensation schemes, moves low ability men towards less competitive schemes, and removes the average gender difference in compensation choices. We also examine between and within-subjects differences in choices for females across the menstrual cycle. We find that the relative reluctance to choose tournaments on the part of women comes mostly from women in the low-hormone phase of their menstrual cycle. Women in the high-hormone phase are substantially more willing to compete than women in the low phase, though still somewhat less willing to compete than men. But the effects of these biological differences can be removed with information as there are no significant differences between the choices of any of these groups after they receive relative performance feedback.

One motivation for experiments on gender differences in competitive behavior comes from the job market, where many careers involve a tournament aspect. An example of this is the corporate ladder, where females make up a small portion of top-level executive positions. Bertrand and Hallock (2001) found that in 1997 the fraction of females in top level management positions was 3% and only 15% of firms had at least one female in a top level executive position. In this paper we explore a potential biological explanation for these differences, and the potential significance of the economic consequences.

There are of course many other explanations for the under-representation of females in higher ranks, such as discrimination and the effect of traditional family roles and raising children on women's career choices and human capital investments Polachek (1981). Jirjahn and Stephan (2004) argue that the attractiveness of piece rate schemes for females is likely caused by the re-

duced possibilities for wage discrimination when performance can easily be measured. It could be for this reason that firms with a higher proportion of females are more likely to offer piece rate compensation Brown (1990). Another possible explanation is a preference by females against competitive, tournament-like situations in favor of alternatives– or a preference by males towards competition and tournaments.

The experimental literature finds that females see lower performance gains from participating in competitive environments Gneezy, Niederle, and Rustichini (2003). Gneezy and Rustichini (2004) find that competition increases the performance of boys, but not girls in athletic running races. Gneezy, Niederle, and Rustichini (2003) also find that in mixed-gender competitive environments males have significant performance increases when an environment is made more competitive, while females do not. However, when females compete only against other females, their performance increases as the environment becomes more competitive. Gupta, Poulsen and Villeval (2005) find that females are more competitive when given the opportunity to choose the gender of a potential competitor. Specifically, females are more likely to choose to enter a tournament if they first choose to be paired against another female before making the tournament entry decision. These results suggest that the gender composition of groups may play a role in performance gains from competition, as well as in the selection into competitive environments.

Overconfidence could be another explanation for gender differences in the willingness to compete, but Niederle and Vesterlund (2007) find that the gender difference remains large even after controlling for the relatively larger overconfidence of men. Dohmen and Falk (2011) report similar results. Gneezy et al. (2009) report results from experiments in a matrilineal society in India where women are more likely to compete than men in contrast to participants in a patriarchal society performing the same type of tasks. Such a result suggests that socialization plays a large role in such gender differences.

Different preferences for risk and/or ambiguity are another possible explanation. In most of the experiments cited above the subjects have very little information about their relative ability when they make their competitive choices. They typically learn their own performance in trial runs, but not the performance of potential competitors. Grossman and Eckel (2008) provide a review of gender differences in risk preferences and find that results generally, but not always, show that females are more risk averse than men. However, most experimental studies examining gender differences for competition argue that gender differences in competitive choices remain after using

various controls for risk preferences. Ambiguity aversion is also a possibility, but ambiguity aversion has not been found to vary systematically across gender. Moore and Eckel (2003) find that females are more ambiguity averse for specific contexts and domains, while Borghans et al. (2009) find that males are initially more ambiguity averse than females, but as ambiguity increases, males and females behave similarly.

None of the results cited so far show that these gender differences in competitive behavior are biologically determined. However, many of the results discuss effects that are potentially endogenous with respect to biological differences such as hormone levels. In this paper, we use systematic variations in the levels of hormones for females across the menstrual cycle to examine a biological explanation for gender differences. We find that women's competitive choices vary substantially across the cycle. Interestingly, the effect is such that during the low-hormone phase of the cycle the behavioral differences are quite large, while in the high-hormone phase females choices to compete are similar to those of males.

Males and females have very different levels of a number of hormones Speroff and Fritz (2005). However, these hormones do not necessarily have the same effects across genders. Most studies examining the effects of hormones have focused on testosterone in males. For males, testosterone levels of financial traders in the morning can predict their daily profits Coates and Herbert (2008). Testosterone levels are correlated with behaviors in economic experiments such as offers and acceptances in ultimatum games Burnham (2007). Financial risk taking has also been linked to circulating levels of testosterone in men (Apicella et al. 2008). Women exhibit large and predictable hormonal variations across the menstrual cycle Speroff and Fritz (2005). For females, estrogen and progesterone have received most of the attention in studies examining behavioral effects.<sup>1</sup>

The mechanisms by which hormones affect behaviors are explained by neuroendocrinological research examining how hormones alter brain activity and emotions. For example, depression and mood have been linked with the density of serotonin binding sites in the brain (Malison et al. 1998). Estrogen affects the density of serotonin binding sites in certain areas of the brain (Fink et al. 1996), including areas of the brain that have been linked to the anticipation and receipt of monetary rewards (Fink et al. 1996, McEwen 2002, Bethea et al. 2002, Platt and Huettel 2008). Progesterone, another hormone that fluctuates across the menstrual cycle, may decrease anxiety and increase sedation Vliet (2001).

---

<sup>1</sup>Estrogen is the generic term for this hormone, estradiol is the form that is most often measured in humans.

It is important to understand that hormones should affect the genders differently because sex differences in the brain develop during perinatal development where both females and male brains are organized differently from different exposure to steroid hormones Gagnidze and Pfaff (2009). Most studies examining hormonal effects have focused on exogenous manipulations in animals. Estrogen in female animals has been found to have masculinizing effects on neural pathways and territorial behavior (Wu et al. 2009).<sup>2</sup> In the adult male brain testosterone acting through androgen receptors is necessary to complement male type behavior Gagnidze and Pfaff (2009). In examining rodents, it has been observed that estrogen is required and received by estrogen receptors to express male-type aggressive and territorial behavior Gagnidze and Pfaff (2009). In a review article, Shepard et al. (2009) conclude that the large literature on sex differences and brain organization indicates that the expression of hormonal effects within a gender may be very dependent on the particular environment under consideration.

Such studies suggest that in human females estrogen may lead to similar behavior as that induced by testosterone in males. Dreher et al. (2007) use fMRI techniques to assess brain activity during the anticipation of uncertain money payments, across different phases of the menstrual cycle. While the study does not include decisions, they find significant changes in activation in areas related to the processing of rewards (such as the striatum) and in the amygdala, an area that activates during fear and anxiety. Two different studies have found behavioral correlations with regard to hormonal fluctuations for females Schultheiss, Wirth, and Stanton (2004); Stanton and Schultheiss (2007). Stanton and Schultheiss (2007) found that estrogen has been found to increase preference for impact or dominance. Given that estrogen and progesterone vary within females, both across the menstrual cycle and depending on hormonal contraceptive use, it becomes important to look at whether the menstrual cycle or hormonal contraceptive use have economically significant effects in labor markets.

Ichino and Moretti (2009) explored the economic effects of the menstrual cycle by using detailed employee attendance data from a large Italian bank. They find that absences for females below the age of 45 tend to occur according to a 28-day menstrual cycle. These 28-day cycle absences explain about one-third of the gender gap in employment absences at the firm. In two experimental studies, Chen et al. (2010) and Pearson and Schipper (2009) explore the possibility that menstrual cycle phase effects drive bidding differences between males and females in auctions. The results

---

<sup>2</sup>This masculinization occurs along the aromatase-expressing neural pathways. Aromatase is an enzyme that converts testosterone to estradiol.

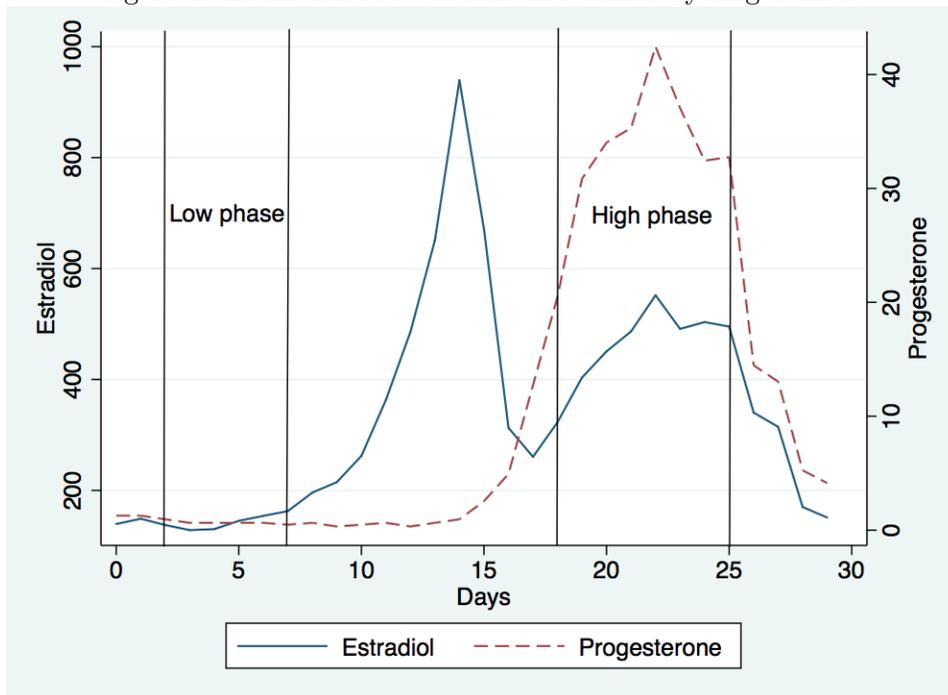
are conflicting, but both studies find that females bid differently depending on their phase of the menstrual cycle. There is one experimental study looking at competitive choices and the menstrual cycle, Buser (2012). This is a between-subjects study of choices of females to compete in all-female groups, and it finds that females participating during predicted high levels of progesterone tend to be less competitive.

Not all studies have found support for hormonal effects on economic decision making. Zethraeus et al. (2009) examine 200 post-menopausal women in a double-blind study. Participants were given either estradiol, testosterone, or a placebo daily for a four week period. No significant differences were found when comparing the behaviors between the three different treatment groups of females a variety of different tasks looking at risk aversion, altruism, reciprocal fairness, trust and trustworthiness. Some research shows that neural receptors in post-menopausal women may have reduced sensitivity to hormonal changes, due to the effects of aging (Chakraborty et al. 2003). Such an aging effect could explain the lack of differences in such a study. Ideally, we would use a double-blind study using exogenous delivered hormones to examine the effects of hormonal differences, but such a study is not feasible with pre-menopausal women, since low hormone levels cause bleeding Speroff and Fritz (2005).

We exploit the variations in estrogen and progesterone levels that occur in females over the menstrual cycle. As shown in Figure 1, both progesterone and estrogen remain low during the early part of the menstrual cycle. This first week of the cycle is when normal cycling females menstruate and can be considered a low-hormone phase. The later part of this is called the pre-follicular phase. Estrogen then rises quickly and spikes just prior to ovulation— this is referred to as the follicular spike. After ovulation (approximately day 14 in the graph), during what is called the luteal phase, females who ovulate experience heightened levels of both progesterone and estrogen. This second spike in both hormones may be referred to as the luteal spike or high-hormone phase (Speroff and Fritz 2005, Stricker et al. 2006). Testosterone levels also vary over the menstrual cycle, peaking just before the follicular estrogen spike (Sinha-Hikim et al. 1998). However, the spike is much smaller than for estrogen, and testosterone levels are insignificantly different during menses and the luteal spike.

Along with comparing female specific differences between a high and low hormone phase, we examine the effects that relative performance feedback may have on both the gender difference to compete and any hormone phase specific differences that occur in low information settings where

Figure 1: Hormonal Fluctuations in Normal Cycling Females



Note: These are median values from Stricker et al. (2006)

individuals do not know their relative rank. In low information settings the effects of gender and menstrual phase are large. A female has a 0.14 lower probability of choosing a tournament compared to a male, even when controlling for performance and confidence. For a female to be as likely to choose a tournament as an average male she must believe she is 40% better than average in performance. We find that the within-gender menstrual phase effect is larger than the across-gender effect. Females in the low-hormone phase of their cycle have a 0.16 lower probability of choosing a tournament than females in high-hormone phase. A low phase female must believe she has 50% better performance to be as likely to compete as a female in the high-hormone phase.

Without feedback, high ability females and males are both more reluctant to enter tournaments than expected value maximization would require. This effect is larger for high ability females. On the other hand, too many low ability types enter competitive environments, and this effect is larger for males. Relative performance feedback moves all these groups toward more optimal choices. This result suggests that the behavioral differences in the willingness to compete are in part driven by stable gender preference differences and specific biological factors such as hormonal variations. But such differences seem to occur because of differing reactions to the generally poor information

concerning a person’s relative rank. Thus, negative effects of biology can be reduced or controlled through the provision of relative performance feedback or through pharmacological manipulations.

## 1 Experimental Design

We use a within-subjects design, scheduling subjects for sessions approximately two weeks apart. An online pre-screening survey was used to recruit and schedule subjects for experiment sessions. We limited the sample of females to those using a monophasic hormonal contraceptive or not using hormonal contraceptives at all.<sup>3</sup> For normally cycling females, the sessions were scheduled during a low-hormone phase (days 2 to 7 in Figure 1) and a high-hormone phase (days 18 to 25 in Figure 1) of the menstrual cycle. These high and low phases are supported by research examining a drop in hormones during menses (Aden et al. 1998). We intentionally avoided the estradiol spike around day 14, because of its short duration and variability within and across females. Other phases were also avoided due to greater measurement error about the hormone changes that could be occurring during those times. This scheduling design has been used in biological research to examine the effects of hormones on the brain Fernandez et al. (2003).

Females using a hormonal contraceptive experience suppression of endogenous hormone production when in the active phase of their contraceptive regimen. Both progesterone and estrogen levels remain fairly constant as the body receives a daily dose of hormones exogenously. During the placebo phase of the contraceptive regimen, there are no exogenous hormones being provided to the body, leading to withdrawal bleeding Speroff and Fritz (2005). We scheduled contraceptive users and normal cycling females accordingly, so that both would be in the experiment during a low-hormone phase and during a high-hormone phase. The high-hormone phase coincides with the luteal spike for normal cycling females and a stable elevated hormone phase for contraceptive users. Since menstrual bleeding is caused by low hormone levels, this allows for easy identification of the low-hormone phase. We avoided the follicular spike, because it is short and difficult to predict and therefore difficult to correctly schedule subjects into sessions.<sup>4</sup>

---

<sup>3</sup>Monophasic hormonal contraceptives release the same level of exogenous hormones each day for the non-placebo phase of the hormonal contraceptive regimen. We excluded users of biphasic and triphasic pills, which provide varying hormone doses.

<sup>4</sup>Ovulation occurs twelve to forty-eight hours after the follicular peak is detected. Ovulation prediction kits capture this rise in hormones; thus, using such kits would not be helpful or cost effective.

Using the screening survey, females were first randomly scheduled during a predicted high or predicted low-hormone phase. A total of 387 females filled out the screening survey which was left open for the entire length of the study. At the time of recruitment, the potential subject pool consisted of a total of 232 females of which 110 participated in experiment sessions. The majority of females who did not participate, but filled out the survey, could not participate due to scheduling conflicts. The low-hormone phase is easily self-identified by the presence of withdrawal bleeding (in both normal cycling females and those in the placebo phase of contraceptives). The high-hormone phase is difficult to identify, particularly for subjects not on contraceptives, because of variability in the cycle. For females using hormonal contraceptives, the high-hormone phase is easier to identify because the hormones become elevated exogenously through the pill or ring. But some hormonal contraceptives, such as biphasic and triphasic pills, vary the levels of hormones on a weekly basis. Based on the screening survey, we found that approximately 15% of pill-users (8% of all females surveyed) take these kinds of pills and these females did not participate in experiment sessions. Instead, we focused on the majority of hormonal contraceptive users, those who use a monophasic form of hormonal contraception, which elevates hormones during the active phase by providing the same dose of hormones exogenously.

Rather than just asking the females who showed up for a session the date of their last or expected next menstruation, we focused on scheduling across two specific hormonal phases to minimize identification error. This was because the menstrual cycle has a large degree of variability, and females may have trouble accurately recalling and predicting menses (Crenin et al. 2004). We also used an exit survey to help minimize errors in classifying phases. Men were simply scheduled for two sessions about 2 weeks apart. We did find that females were less likely to attend experiment sessions during the low phase of the cycle; thus a potential bias may exist in experimental settings.<sup>5</sup>

Previous studies on differences in competitive choices have used between-subjects designs (Niederle and Vesterlund 2007, Gneezy et al. 2009). In our within-subjects design each subject participated in one session of math tasks, and another of word tasks. We used two different tasks in part because we wanted to minimize behavioral spillovers from the first to the second session, and in part because it is generally believed that females may view themselves as having worse math skills than males Niederle and Vesterlund (2010). For this reason females may be less likely to compete in math

---

<sup>5</sup>A discussion of this attendance bias is available from the authors.

tasks than in word tasks. This design is the first to examine whether there are stable differences in competitive choices across genders between-subjects and within-subjects for math and word tasks.

Subjects were randomly assigned to start with a math or a word based session. In each session tasks were performed for five different treatments, one of which was randomly chosen for payment at the end of the experiment. Each treatment lasted 4 minutes. In the first treatment participants performed the task under a non-competitive piece rate compensation scheme, where pay was entirely dependent on the individual's own performance. In the second treatment, participants were randomly assigned to a winner-take-all tournament with a size of two, four, or six competitors. This second treatment provided participants with experience in a situation where their pay depended on their own performance as well as the performance of others. In the third treatment, participants performed the task with a group pay (revenue sharing) form of compensation. This treatment randomly paired participants and payment for the group's total production was split evenly. This third treatment can be considered the least competitive because of the possibility of freeriding. It can be shown that given some random assignment of competitors or group members, this design should lead low ability individuals to choose group pay and high ability individuals to choose a tournament.<sup>6</sup>

In the final two treatments subjects were able to choose between piece rate, group pay, or a two, four, or six person tournament. Before the fourth treatment, subjects were told their own absolute performance from treatment 1, but were not told anything about the performance of others. Just before the fifth treatment, participants were shown how all individuals in the session had performed in the first treatment with their performance highlighted, and they then chose their compensation method and performed the task again.

The math task was similar to the one used in Niederle and Vesterlund (2007). Participants were asked to add four randomized two-digit numbers and complete as many of these summations as possible in 4 minutes. Equations were presented to participants on a computer screen and they typed in their answer and pressed the *Enter* key or clicked a *Submit* button on the screen. After each submission participants were promptly shown the next equation to solve, using scratch paper if they wanted. On the screen, the equations looked like the following:

$$12 + 57 + 48 + 52 =$$

---

<sup>6</sup>A model justifying such predictions is provided by the authors in an online Appendix.

The word task was similar to that used by Günther et al. (2010). In this task participants are shown a letter on a computer screen and have four minutes to form as many unique words as possible that begin with that specific letter. The letter remains on the screen for the entire four minutes and participants enter in their word submissions in a text box below the letter. The attempted word formations are then listed below the text box to help subjects minimize duplicate answers, since these are counted as incorrect. Common place names (cities, countries) are acceptable, but proper names are counted as incorrect. Plural and tense changes to root words are counted as separate and correct answers as long as these words still begin with the appropriate letter. In the experiment, participants were informed of the rules before beginning the task.

All participants were informed that everybody in the same session and same treatment received the same letter, to ensure a task of equivalent difficulty for all participants in each treatment. The word list used for grading words is a common English word list used by open source word processors.<sup>7</sup> We used a restricted group of letters for this study to limit the variation of difficulty between treatments and sessions (e,f,g,h,i,l,n,o). Between 2.7% to 3.8% of all words in the word list began with these letters.

For the piece rate compensation, the payoff an individual receives is equal to the piece rate multiplied by the production of the individual for that particular treatment. Payoffs for both the math and verbal tasks were calculated in a similar manner though the base rate was different for word formation tasks (\$0.25) and math addition tasks (\$0.50), to adjust for generally higher performance in the word task. In a tournament, if an individual has the best performance in his group then he receives the piece rate multiplied by the size of the tournament, multiplied by his individual performance. If an individual does not have the best performance in his tournament then he receives nothing. In the event of a tie, the individual receives a fraction of the tournament winnings based on the number of individuals he tied with. Subjects were not informed about whether they won or lost a tournament until all five treatments were complete. After each treatment, and before seeing their score, subjects were asked how well they thought they did and how well they thought the average person in the session did, and they were paid for having accurate predictions.

Subjects were told that they could be randomly grouped with people that did not necessarily choose the same compensation option and that they therefore could be playing under different rules than their potential competitors or group members. This strengthens the incentive for high ability

---

<sup>7</sup>Spell Checking Oriented Word Lists (SCOWL), Revision 6, August 10, 2004 by Kevin Atkinson.

types to choose a more competitive tournament, since there is a positive probability that they may compete against lower ability individuals. This rule also creates an incentive for low ability individuals to choose group pay, as they may be matched with high ability individuals who would increase their expected payoffs.<sup>8</sup>

## 2 Results

Experiment sessions took place in a computer lab at a large public university and all IRB procedures were followed. The majority of the 219 participants were university students, characteristics are shown in Table 1. The average size of the 26 sessions was 13.31 participants (with a standard deviation of 4.05). Sixty-two female and 64 male subjects participated in both sessions. Using the pre and post surveys we were able to classify 45 females as participating in a session during a low-hormone phase of their menstrual cycle, and 34 during both a low and a high-hormone phase. The word task was used in 12 of the sessions and the math task was used in 14 sessions. Of the 345 individual subject sessions, 165 involved the use of the word task and 180 used the math task.<sup>9</sup>

Table 1: Summary statistics of session attendees

Participant Characteristics					
Variable	Mean	Std. Dev.	Min.	Max.	N
Age	20.52	2.81	18	33	218
Years PS	2.18	1.48	0	6	217
GPA	3.29	0.47	2	4.1	218
Live Independently	0.82	0.39	0	1	219
Female	0.5	0.5	0	1	219
Meds	0.09	0.28	0	1	219
Characteristics by Sessions*					
Low Phase	0.14	0.34	0	1	345
Word task	0.48	0.5	0	1	345
Second session	0.37	0.48	0	1	345
Session Size	13.26	4.14	7	21	26

Note: 126 individuals attended a second session.

Table 2 shows that men and women were similar in terms of age, GPA, years of post secondary schooling (Years PS) and even have the same proportion taking psychological medication (Meds). Both genders were assigned to sessions with similar characteristics, except that on average females

<sup>8</sup>The text for experiment instructions is provided by the authors in an online Appendix.

<sup>9</sup>There were a total of 346 subject sessions, but one female was removed from the data due to non-compliance with the task instructions.

were in slightly larger sessions. The proportion of females per session ranged from 0.25 to 0.7 and averaged 0.5. Thus, all sessions had some degree of gender mix and on average this mix was about one-to-one.

Table 2: Mean values of individual and session characteristics by gender.

Sex	Age	Years PS	GPA	Indep.	Meds	Word	Size	Sess. 2
Male	20.70	2.14	3.25	0.83	0.09	0.47	13.79	0.37
Female	20.35	2.21	3.33	0.80	0.08	0.48	15.30	0.36
Total	20.52	2.18	3.29	0.82	0.09	0.48	14.54	0.37

Sessions took place three to four times a week and were held in the morning. Each session took slightly less than an hour, including approximately 10 minutes at the beginning of the session during which participants waited together in a foyer. This allowed participants to see that sessions included both males and females. Once participants entered the lab, partitions were used so that participants could not see each other’s computer screens or facial responses from the feedback received. Competition and group memberships were also anonymous.

Payouts were based on one randomly chosen treatment, excluding the flat rate show-up payment, payouts averaged \$7.38 for the math session and \$15.01 for the word sessions. Participants who attended two sessions were later asked to perform a risk aversion task similar to that used in Holt and Laury (2002). The risk aversion tasks were performed a few days after the second session to avoid endogeneity with competition task earnings. A total of 112 participants (56 male and 56 females) participated in the risk aversion task. The average payout for the risk aversion task was \$6.57.

## 2.1 Task Performance

Each individual participated in five different treatments in each session. For the first three treatments the compensation schemes were as follows:

Treatment 1: Piece rate (\$0.50 per sum and \$0.25 per word).

Treatment 2: Random sized tournament of 2, 4, or 6 individuals (the winner earned the piece rate multiplied by the size of tournament).

Treatment 3: Group pay: an individual was paired with a randomly chosen partner and the total production of the 2 individuals was multiplied by the piece rate and then split evenly.

Table 3: Performance Across Task Type, Treatments and Gender

Math	T1	T2	T3	T4	T5	Word	T1	T2	T3	T4	T5
Male	10.9	12.1	12.3	12.7	12.8	Male	38.2	39.4	43.0	45.3	47.0
Female	9.9	11.4	11.8	12.3	12.1	Female	41.0	41.1	45.0	48.4	47.3
Both	10.4	11.8	12.0	12.5	12.5	Both	39.6	40.3	44.0	46.9	47.1

Note: The column notation refers to treatment order.

For example, T1 refers to the first treatment.

Table 3 shows mean performance by gender over treatments and tasks. The increasing mean values over the first three treatments in both the math and the word tasks suggest that subjects are learning to do the task better during the session. There are no statistically significant performance differences between males and females in either task.<sup>10</sup> This lack of a performance difference by gender, for either task, removes one obvious potential reason for gender differences in choices.

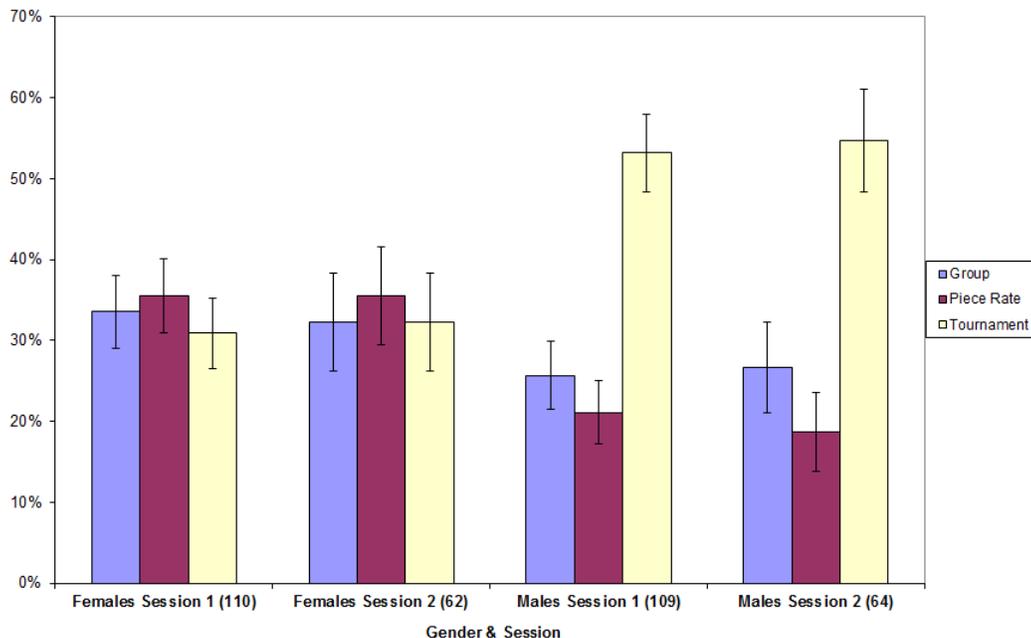
## 2.2 Gender Differences in Competitive Choices

Niederle and Vesterlund (2007) and Gupta et al. (2005) find that when given the choice between a tournament and piece rate females are less likely than males to enter tournaments. To test whether this basic result can be replicated with our protocol, we focus on choices made in Treatment 4. In those studies, individuals did not have information about their relative performance, and in our study this feedback comes only after Treatment 4. The available choices of group pay, piece rate, and tournaments of increasing size can be ordered by increasing competitiveness, with sharing being less competitive than not sharing and larger tournaments being more competitive. In the figures and empirical analysis we lump the two, four and six person sized tournaments together though the results are robust when using an ordered scale for tournament size.

Figure 2 shows the distribution of choices made by males and females in the first and second sessions for Treatment 4. The gender differences are large: pooling over sessions we find that only 31% of females chose to compete in tournaments while 54% of males chose the tournaments. The difference persists for the piece rate: 36% of females chose the piece rate compared to only 20% of males. These differences are all significant at the 2% level or better with chi-square tests. This replicates earlier findings, and shows that gender differences for competitive choices are robust to the addition of a group pay option and different sized tournaments. We also find that, on average,

<sup>10</sup>An analysis of the performance effects that occur from learning and different competitive settings are available on request.

Figure 2: Choice Differences Before Feedback, by Gender and Session



Note: Sample size in parentheses.

males and females chose consistently across the two repeated sessions, despite the fact that these which are often separated by weeks.

While there are no significant differences in performance between males and females, other factors such as age and GPA might conceivably affect compensation choices. In this design, we predict that with full information about abilities individuals would sort according to ability, with the least able individuals choosing the least competitive environments and the higher ability individuals choosing more competitive tournaments. Given this we use an ordered probit to test whether the gender differences in the probability of selections remain after controlling for other potentially relevant factors.

Table 4 shows that the gender differences persist with these controls, along with the addition of control variables for confidence, performance, and improvement in the repetition of tasks in a tournament. Columns 1 to 3 use CompScale as the dependent ordinal variable, where group pay compensation is less competitive than piece rate which is less competitive than a tournament of any size.<sup>11</sup> In the results, we include both random effects estimations using an ordered probit

<sup>11</sup>Our results are consistent with a multinomial logit model and from using ordered probits with rankings that treat larger tournaments as more competitive.

model as well as pooled models with clustered standard errors based on the individual and the session. For nonlinear estimations such as ordered probits, random effects models are often used to deal with the difficulties and bias involved with using fixed effects models Arellano and Honoré (2001). Given that the experiment data is considered a short panel, any fixed effects estimation of a nonlinear model would also suffer from the well-known incidental parameter problem that may bias fixed effects results Greene (2004). For these reasons we chose to use a random effects ordered probit for estimation purposes and also compared results to ordered probit models that cluster the standard errors by individual or by experiment session, we find our results to be consistent using all these methods.

Table 4: Ordered Probit: Choices for No Relative Information (Treatment 4)

	(1)	(2)	(3)	(4)
Female	-0.40*** (0.15)	-0.49** (0.19)	-0.35*** (0.135)	-0.35*** (0.114)
Confidence (T1)	0.98*** (0.29)	0.99*** (0.34)	0.87*** (0.256)	0.87*** (0.246)
Improve (T2)	0.72*** (0.23)	0.73** (0.32)	0.62*** (0.181)	0.62*** (0.187)
%-tile Rank (T1)	1.08*** (0.26)	0.85*** (0.32)	1.04*** (0.227)	1.04*** (0.233)
Clustering	RE	RE	Individual	Session
Clusters			220	26
Demographic Controls	Yes	Yes	Yes	Yes
Risk Controls	No	Yes	No	No
Observations	343	224	343	343
ll	-335.6	-212.3	-336.5	-336.5
chi2	61.00	48.81	71.85	116.8

Note: RE means that individual random effects were used. ll refers to log-likelihood. Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.10

Table 4 replicates the results of Niederle and Vesterlund (2007) with Treatment 4, before relative performance feedback. Females are less likely than males to enter tournaments, even when controlling for individual confidence (Confidence (T1)) and relative rank of performance within the session (%-tile Rank (T1)) from the first treatment. The %-tile Rank (T1) variable gives the rank of an individual based on her or his performance in Treatment 1 in the session. Using rank allows us to have the same measure for both math and word tasks.<sup>12</sup> Confidence is measured by an individual's predicted performance at the end of Treatment 1 (prior to finding out their actual performance)

<sup>12</sup>Using a variable that measures actual performance with an interaction term for the type of task, gives the same results as are presented here.

divided by that individual’s prediction of the average performance of all session participants.<sup>13</sup> To control for performance, we use the relative rank from Treatment 1, but the results are unchanged when using absolute performance along with an interaction term for word based tasks.

As expected, both confidence and the actual percentile rank from the first treatment are positively correlated with the selection of more competitive environments. Improvement in performance between the first and second task (Improve (T2)) also has a significant positive effect. These regressions include controls for individual specific characteristics, including the number of years of college, psychoactive medication, GPA, and age.<sup>14</sup> Column 2 includes a measure of risk aversion for individuals that participated in a task similar to the one used by Holt and Laury (2003). We find that this measure of risk aversion is not significantly correlated with competitive choices in Treatment 4.

In calculating the marginal effects we find that a female has a 0.14 lower probability of choosing a tournament than a male, even when controlling for performance and confidence. For a female to be as likely to choose a tournament as an average male, we would have to increase her belief about her performance relative to the average by 40%, which is a significant increase in overconfidence. A ten-percentile improvement in actual relative performance would increase the probability of entering a tournament by 0.04. A female would have to improve her percentile rank by 34% to be as likely to enter a tournament as a male. Thus, these gender differences are not just significant, but they are also large.

After each treatment, before receiving any feedback, subjects were asked how many correct answers they believed they submitted. Subjects were paid (\$0.25) for each correct answer to encourage accurate answers. We create a measure of confidence by dividing an individual’s prediction of how well he or she did divided by his or her prediction of the session average for that treatment. Since the average individual should believe they did not perform any better than the session average, this confidence measure should have a mean of one— in the absence of overconfidence.<sup>15</sup> We could have asked for rank estimates instead of performance estimates, but rank is a poor measure

---

<sup>13</sup>Females tend to be less confident than males in the math task, but females and males have no significant differences in confidence in the word task and selection differences still remain.

<sup>14</sup>Details on these controls are provided by the authors in an online Appendix.

<sup>15</sup>We also asked how many correct answers they believed were submitted by the most productive person, the least productive person, as well as the average number of correct answers, for each session and treatment. We use the average instead of the prediction of the best or worst individual in the session because it provides a clean measure of overconfidence. In other estimations, not included here, these measures were separated and variations of using both the best performance and the worst performance as the denominator were used with little difference in our results.

of the degree of over or under confidence. Consider two individuals that think they are ranked first in their respective group. One may think that he is 10% better than the average while the other may think she is 50% better. Both these individuals would be treated as having the same level of confidence with the rank measure, but one individual is actually much more confident. We use the measure of confidence from the first treatment because every subject performed the task for this treatment under the same piece rate form of compensation. This confidence variable provides the earliest measure of overconfidence before experiencing any feedback or differing experimental manipulations.

Changes in performance as the experiment proceeds could also change confidence. The variable Improve (T2), measured as the ratio of the individual's performance from Treatment 2 divided by the performance in Treatment 1, captures the effect of individual improvement between Treatment 1 (piece rate) and Treatment 2 (tournament). There are two possible reasons that this variable should matter: First, individuals may feel that they improve more than the average individual or that they were unlucky in Treatment 1 compared to how others would have performed. Second, it may be the case that individuals become more motivated to put in greater effort because of the competitive nature of the tournament in Treatment 2. Individuals that improve a lot from competing in such settings would be more likely to choose to compete than individuals whose performances are not positively affected by competitive settings.

Niederle and Vesterlund (2007) found that part of the difference between male and female willingness to compete was driven by males being more overconfident than females. In their study, independent of confidence, females had a 0.16 lower probability of entering a tournament than males. Using our measure of confidence we find that the gender difference is nearly the same, 0.14.

Since we have multiple observations from the same individuals and individuals participate in the same sessions, we also run regressions where we cluster standard errors on experiment sessions and then also separately on individuals. The final two columns in Table 4 show that the results concerning females being less likely to enter in tournaments without relative performance feedback remain consistent when using errors that are clustered on the specific experiment session or on the individual. In this table the dependent variable is the same ordered variable of competitiveness used previously where group pay is less competitive than piece rate and piece rate is less competitive than a tournament of any size.

Our within-subjects design includes one session of math treatments and one of word treatments. Günther et al. (2010) found that in a maze task, men increased performance in reaction to competitive pressures by more than women did. In a word task the improvements were the same. They attribute this to a "stereotype threat" arising from beliefs that women are not good at the maze task. This could logically lead to different choices by women to compete, with different tasks. We find there is little difference in the selection of competitive environments by females regardless of the type of task used. In both math and word tasks females are less likely to compete in the treatment before receiving performance feedback, approximately 30% of females enter tournaments in both tasks. We also find little difference in choices by males as more than 50% of males chose to compete in tournaments in both math and word tasks.

Table 5 looks at confidence differences by gender and task. Both genders are overconfident on average. Males are significantly more overconfident in their math abilities than females, and there is no significant difference in confidence between males and females in the word task. There is no significant difference among females between the math and word tasks, while males are significantly more confident in their math performance than in their word task performance. On average, males are slightly more confident in their abilities than females. This is partly driven by a few high ability males who are correct in believing they are better than the average, but overestimate the degree. For example, the highest level of confidence for a male is 3.38 times his prediction of the average. His actual performance is 2.29 times the actual session average. Overall, males and females are fairly consistent in their choices to compete in both types of task: males choose to compete more than females in both math and word tasks even though male overconfidence is higher in the math task. The type of task was not significant in regressions for choices, with or without confidence controls.

Table 5: Confidence Differences by Gender and Task Type

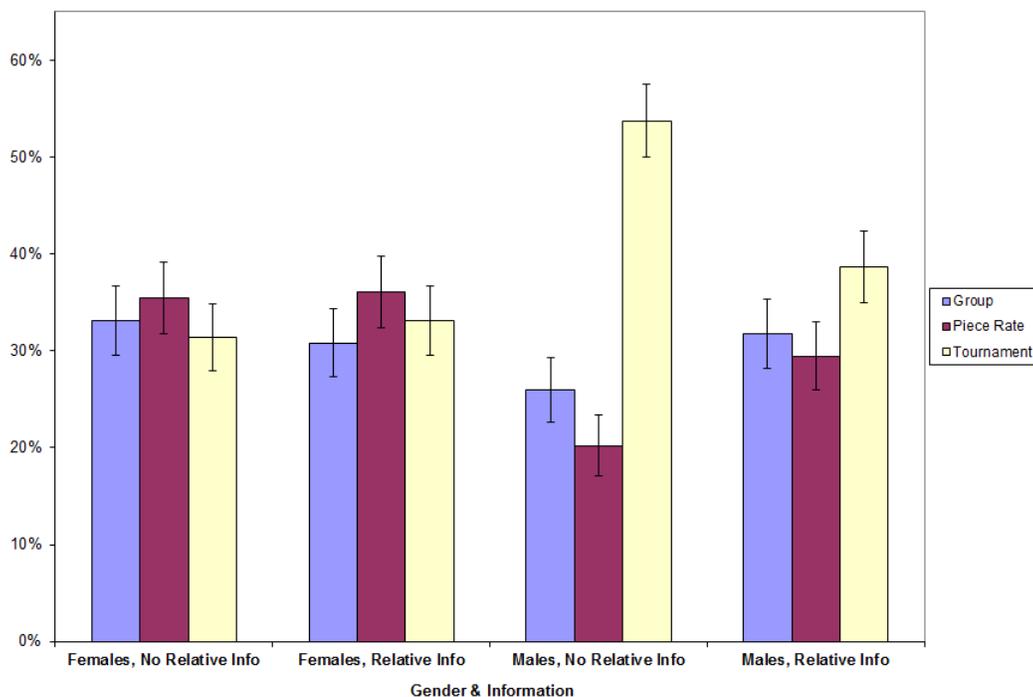
Task Type	Gender	Confidence (T1)	S.E.
Math	Female	1.081	0.030
Math	Male	1.209	0.037
Math	Both	1.145	0.024
Word	Female	1.046	0.029
Word	Male	1.039	0.030
Word	Both	1.043	0.021

Note: T1 refers to the first treatment.

### 2.3 Performance Feedback Eliminates Gender Differences to Compete

Providing information about the quality of possible competitors might reduce mistakes in competitive choices, but there is no obvious reason feedback should reduce the gender difference in choices, if that difference is primarily driven by preferences. We test the effect of performance feedback on choices by providing subjects with an ordered list of the performance of all the participants in their session from Treatment 1, with their own performance highlighted, before they choose their Treatment 5 compensation scheme. This provides information about the quality of their potential competitors, if they choose to enter a tournament.

Figure 3: Choice Differences by Gender and Information Treatment (Treatments 4 and 5)



Note: Females (172), Males (173). Sample size in parentheses.

The two groups of bars on the left side of Figure 3 suggest that females' choices are barely affected by information about the performance of potential competitors. The right side of the figure shows that males' choices change dramatically. There is a significant increase in the proportion of males choosing piece rate (5% significance level) and group pay (10% significance level), and a significant decrease in the proportion choosing tournaments (5% significance level). Comparing the distributions of men's and women's choices in Treatment 4 gives a Pearson chi-square statistic

of 18.79 (p-value: 0.000). After relative performance feedback in Treatment 5 male and female competitive choices are not significantly different (chi-square statistic is 1.91, p-value: 0.385).

Table 6 shows the results from two different types of ordered probits for Treatment 5 choices, using the CompScale competitiveness definition from the Treatment 4 analysis. Columns 1 and 2 show, that once performance feedback is provided, there are no significant differences between male and female choices. Instead, we find that choices are very dependent on the relative performance information, and on the individual’s improvement from Treatment 1 to Treatment 2. Risk aversion control variables are not significantly correlated with compensation choices on average; though risk aversion measures were significant when only examining high ability individuals’ choices in Treatment 5. The one variable that consistently affects individual choices in Treatment 5 is an individual’s percentile rank from Treatment 1, a summary statistic of the feedback information provided before the Treatment 5 choice.

Table 6: Ordered Probit: Choices After Feedback (Treatment 5)

	(1)	(2)
Female	-0.02	0.13
	(0.18)	(0.21)
Confidence (T1)	0.44	0.65*
	(0.30)	(0.35)
Improve (T2)	1.01***	0.65**
	(0.26)	(0.32)
%-tile Rank (T1)	2.59***	2.31***
	(0.34)	(0.37)
Risk Controls	No	Yes
Demographic Controls	Yes	Yes
Observations	343	224
ll	-316.7	-194.5
chi2	98.51	79.67

Note: Estimations include individual random effects.

ll refers to log-likelihood.

Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

The overall conclusion from Figure 3 and the probit models in Table 6 is that there are no significant gender differences in competitive choices when subjects are fully informed of their relative performance compared to potential competitors. In the next section, we consider the costs of the selection differences between men and women when they lack information about the quality of competitors and whether there are gender differences according to ability levels.

## 2.4 The Cost of Choices and the Effect of Feedback, by Gender and Ability.

To give some sense of the costs of gender differences in choices, we simplify and assume people maximize expected payoffs, keep effort constant across compensation choices, and take the choices and performance of others as given. Table 7 shows the average expected value losses for the suboptimal selections by males and females in Treatment 4 and Treatment 5.<sup>16</sup> Each column represents the optimal choice that should have been made. The numbers represent the average expected value cost for choosing something other than that optimal choice. For example, in the first row under column 6 (for the 6 person sized tournament), the 27.27 represents the average loss to females whose optimal choice was a tournament of six, but who instead chose a different form of compensation. The Avg Loss column provides the average loss by gender and treatment. The average loss of 6.78 in the first row means that females lost an average of \$6.78 from their suboptimal choices in Treatment 4.

Table 7: Average Loss (\$) from Suboptimal Choices

Treatment	Gender	Optimal Choice					Avg Loss
		Grp	PR	2	4	6	
4	Female Avg Loss	1.58	2.18	2.91	6.80	27.27	6.62
4	Male Avg Loss	2.42	3.09	2.31	3.29	12.60	4.89
5	Female Avg Loss	0.88	1.83	2.21	5.93	18.70	4.70
5	Male Avg Loss	1.39	1.44	2.02	4.79	10.98	3.88

Note: Grp refers to group pay. PR refers to piece rate.

The numbers of 2,4,6 refer to the tournament size.

In Treatment 4, the average expected value loss from selection mistakes was \$4.89 for males and \$6.62 for females, a statistically insignificant difference with a t-test. These loss differences are mostly driven by high ability females choosing not to compete, and to a lesser extent by low ability males choosing to compete. Column 6 shows that many high ability females (those who should select a tournament size of 6) are instead selecting smaller tournaments or group pay or piece rate, at a large cost. The top females lose \$27.27 in expected value compared to \$12.60 for the top males. In contrast, low ability males make only slightly more costly decisions than low ability females, averaging \$2.42 versus \$1.58 for the lowest types of each gender. We find that high ability females and high ability males are not entering competitive environments enough. But the high ability females overwhelmingly select the noncompetitive environments of piece rate and group pay,

<sup>16</sup>The method used to calculate expected values is based on using the percentile rank as the probability of success within a tournament.

which are very costly decisions. In contrast, too many low ability males are entering competitive environments, but these mistakes are not particularly costly, on average, because low ability males would not perform well in the piece rate either.

Table 7 also shows that relative performance feedback decreases the average expected value losses for both males and females and shrinks the gender gap as well. The decreases in expected value losses are greatest for high ability females, whose average expected loss fell from \$27.27 in Treatment 4 to \$18.70 in Treatment 5, while losses for high ability males fell from \$12.60 to \$10.98. Low ability females and males tend to move towards group pay as they get performance feedback. While a gender difference remains, with low ability males making more expensive mistakes than women, the cost differences are small.

We now consider the question of how relative feedback information affects the choices of high ability females and males. A high ability individual is defined as an individual who should enter a four person tournament or larger to maximize expected returns from competition. We find that relative performance information leads to a large increase in the proportion of high ability females entering tournaments. Over 50% of high ability females enter tournaments when given relative performance feedback, which is significantly more than the 31% that choose tournaments before receiving the performance feedback. In testing for distributional changes, we find that there is a significant difference in choices for females between Treatment 4 and Treatment 5; using a Pearson chi-square test the level of significance is  $p = 0.034$ .

With information, fewer high ability males enter tournaments (12% fewer), but this change in tournament selection is not statistically significant at the 5% level. The distributional difference of choices for high ability males coming from information feedback is not significant as a chi-square test comparing high ability males between treatments produces a level of significance of  $p = 0.317$ . Without feedback in Treatment 4, there is a significant difference in the distributions of competitive choices between males and females ( $p = 0.000$ ). After receiving feedback, the distributions exhibit smaller differences as the level of significance using a chi-square test is  $p = 0.158$ . Thus, relative performance feedback seems to eliminate most of the differences in choices between high ability females and high ability males.

Given that high ability individuals exhibit behavioral changes from relative performance feedback then it is also worth considering how this information affects low ability individuals' choices to compete. We consider a low ability individual to be an individual with performance below the

median in her or his respective session from Treatment 1. The largest effects for low ability types are for males. Information drops the percentage of low ability males choosing tournaments from 43% to 22% and increases the percentage of low ability males choosing group pay from 37% to 51%. For low ability males, the difference in the distribution of competitive choices between Treatment 4 and Treatment 5 is significant at a  $p = 0.010$  using a chi-square test. No such significant difference occurs for low ability females. We find that with no performance feedback in Treatment 4, the distributions of choices are significantly different between low ability females and low ability males as a chi-square test leads to a  $p = 0.054$ . But in Treatment 5, after participants receive performance feedback, there are no significant differences between distributions for low ability females and males.

Information about relative performance moves high ability females towards more competitive choices and low ability males away from tournaments towards less competitive types of pay. Low ability females show only a small movement away from group pay towards piece rate. Overall, providing relative performance feedback information leads to more efficient sorting by both genders.

## 2.5 Competitiveness Differs Between High and Low Hormone Phases

Normal cycling women experience large changes in hormone levels across the menstrual cycle (Figure 1). Women using hormonal contraceptives experience hormonal variations between the placebo phase of the contraceptive regimen and the active phase (when pills with hormones are being taken or a ring is being used), these exogenous variations are similar to the High phase and Low phase shown in Figure 1. Table 8 summarizes the responses of females to the screening survey used to schedule females. Of the females who completed the screening survey almost 15% missed a menstrual period during the previous 3 months. Over 43% of these females experienced menstrual cycle irregularity of 3 days or more, suggesting that predicted menstrual periods may have significant measurement error. Due to the potential inaccuracies introduced by this prospective survey we also used an exit survey with both retrospective and prospective questions on menstruation to classify hormonal phases for our analysis.<sup>17</sup>

---

<sup>17</sup>Missed periods are a problem for identification purposes in normal cycling females as they imply that a female may not have ovulated during that month, and thus did not experience a mid-luteal peak in hormones. Furthermore, without a recent menstrual period it is difficult to determine the current phase in the hormonal cycle.

Table 8: Menstrual Cycle Regularity

Regularity of Period	Percent	Count
Identical	14.3%	55
Within 1-2 days	42.3%	163
Within 3-7 days	34.3%	132
Very Irregular (7+)	9.1%	35
Total		385
Missed Period in Last 3 Months	Percent	Count
Yes	14.7%	57
No	85.3%	330
Total		387

Note: Numbers may not add up due to item non-response in screening survey.

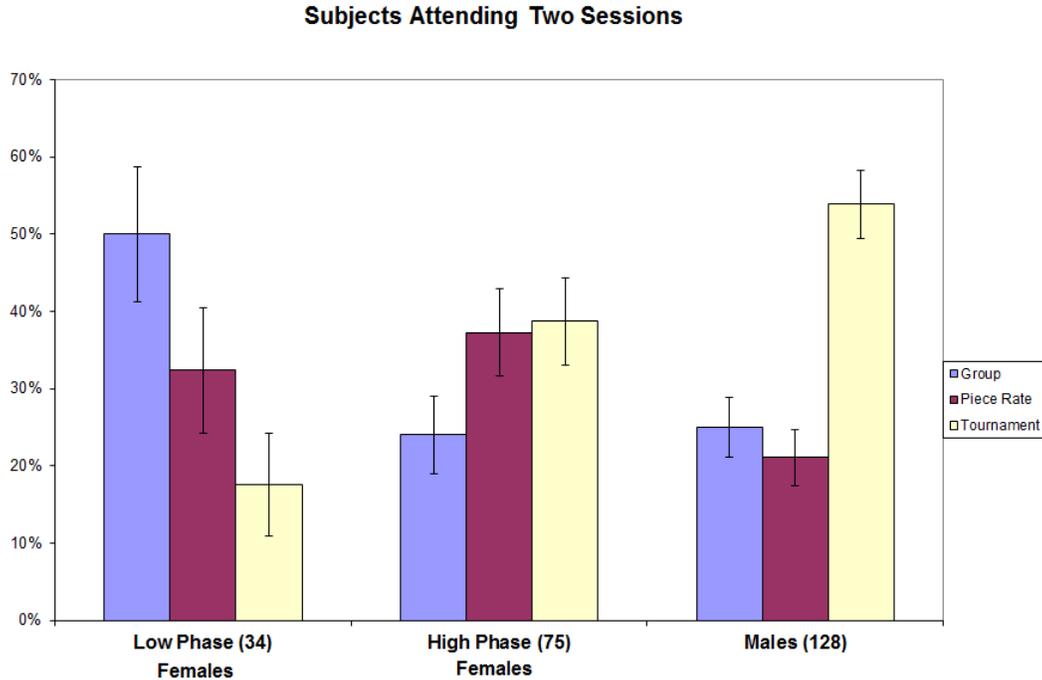
The screening survey also provided information on the proportion of females that use hormonal contraceptives. Over 54% of females in our sample used some form of hormonal contraceptive in the form of the pill or ring. This makes for easier predictability of low and high phases for these females, since hormonal fluctuations are exogenously determined by hormonal contraceptive use. To help identify hormonal phases for females using a hormonal contraceptive, we asked all female participants for the start day of their hormonal contraceptive regimen.

Of the females that participated in experiment sessions, 62.7% of those attending a first session were following a hormonal contraceptive regimen, as were 62.9% of those at second sessions. The American College Health Association found that about 72% of sexually active females were using some form of hormonal contraceptive in 2008. In examining contraceptive use by females in the United States, it was found that for women between the ages of 15 to 44, over 82% had at one time taken oral hormonal contraceptives (Mosher et al. 2004), suggesting that our sample is not unusual in terms of contraceptive use.

We hypothesize that the low-hormone phase, whether induced through endogenous or exogenous means, is associated with similar behavioral changes for both hormonal contraceptive users and normal cycling females. We tested this by controlling for hormonal contraceptive use and found no systematic significant difference in behavior between hormonal contraceptive users and normal cycling females. We therefore pool both groups of females and focus on similar differences across the two hormonal phases.

Figure 4 shows the distribution of competitive choices of females by phase along with choices by males, in treatment 4 before participants had relative performance feedback. Female behavior

Figure 4: Competitive Choice by Gender and Hormonal Phase



Note: Sample size in parentheses.

is very different across the two phases. Females are more than twice as likely to choose group pay when they are in the low phase, and twice as likely to choose tournament when they are in the high phase, though still not as likely as men. When we include controls in regressions, this last difference will become insignificant. The data for the histogram includes all females and males that attended two sessions and all females who could be identified as being in the low or the high phase. Due to the difficulty of predicting the low phase, some females were identified by the exit survey as being in the same phase for both word and math tasks. As well, some phases could not be accurately identified and those subjects are not included in the analyses.

These differences in competitive environment choices across hormonal phases may result from differences in expected performance changes across the menstrual cycle, or from different preferences for competition. We find that for the most part, there are no significant performance differences between females in the low phase and those that are not in the low phase.<sup>18</sup> It is also possible that females in a specific hormonal phase might experience greater aversion to certain types of tasks;

<sup>18</sup>These performance differences are provided by the authors in an online Appendix.

therefore, we also examined choices separately for math and word tasks. Females that participated in a math or word task during the low phase were then scheduled for the other type of task when in a high phase, and vice-versa. We found that the general correlation between competitive choice and menstrual phase holds across tasks: high phase females are less likely to choose group pay and more likely to choose tournaments in both word and math tasks. For both math and word tasks, the majority of low phase females avoided tournaments and chose group pay, while females in the high phase chose tournaments significantly more while choosing group pay the least out of the three choices.

Table 9 uses the CompScale variable as the dependent variable, it is an ordered categorical variable with choices ranked from group pay to piece rate to tournament. The first to third columns provide estimates using random effects ordered probit models while the final two columns use pooled ordered probit models with standard errors clustered by individual and by session respectively. Almost all the columns, with the exception being the second column, include both males and females. The second column consists of a female only sample and the third column takes into account only males and females for which risk aversion measures were available and could be used as control variables.<sup>19</sup>

We find that females in the low phase select noticeably less competitive compensation plans than females in the high-hormone phase. In fact much of the average difference in competitive choices between males and females is driven by the choices of the low phase females. This result holds even when controlling for confidence. It is worth noting that there are no significant differences in confidence levels between low hormonal phase and high hormonal phase females, and yet females in the low phase avoid the competitive environments of tournaments and are more likely to choose the least competitive setting possible, group pay.

These choice differences could potentially result from discomfort during the low-hormone phase of menstruation. But females in the low-hormone phase do not behave differently from any other group once they receive relative performance feedback. Thus, physical discomfort is an unlikely explanation for these systematic differences in low information settings.

The magnitudes of the marginal effects (calculated using the pooled cross sectional estimates) of being in the low-hormone phase are substantial and are larger than the average gender effects. For group pay, females on average have a 0.08 higher probability of choosing group pay than males.

---

<sup>19</sup>All regressions include controls for session ordering, GPA, age, education, and psychoactive medications.

Table 9: Ordered Probit: Hormone Effects for No Relative Information (Treatment 4)

	(1)	(2)	(3)	(4)	(5)
Sample	All	Females	Risk	All	All
Female	-0.29*		-0.26	-0.25*	-0.25*
	(0.16)		(0.21)	(0.149)	(0.137)
Low Phase	-0.46**	-0.53**	-0.76***	-0.45**	-0.45***
	(0.22)	(0.26)	(0.27)	(0.200)	(0.156)
Confidence (T1)	0.91***	1.08**	0.90**	0.84***	0.84***
	(0.30)	(0.54)	(0.35)	(0.275)	(0.234)
Improve (T2)	0.69***	0.79**	0.72**	0.62***	0.62***
	(0.23)	(0.38)	(0.32)	(0.182)	(0.187)
%-tile Rank (T1)	0.99***	0.52	0.72**	0.96***	0.96***
	(0.26)	(0.43)	(0.32)	(0.225)	(0.227)
Clustered	RE	RE	RE	Individual	Session
Clusters				215	26
Risk Controls	No	No	Yes	No	No
Demographic Controls	Yes	Yes	Yes	Yes	Yes
Observations	328	155	211	328	328
ll	-321.7	-156.0	-197.4	-322.2	-322.2
chi2	58.60	19.76	51.31	69.61	124.6

Note: 45 low phase females could be identified for data analysis.

RE means that random effects were used. ll refers to log-likelihood.

Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Females in the low phase have an additional 0.16 higher probability of choosing group pay. For tournaments, females have a 0.10 lower probability of choosing a tournament when compared to males, and females in the low-hormone phase have an additional 0.16 decrease in the probability of choosing a tournament.

The changes over the menstrual cycle are also large relative to the effects of confidence and performance. For a female in the low phase to have the same probability of entering a tournament as a female in the high phase we would have to increase her belief about her performance relative to the average by 50%. In terms of an equivalent performance effect, a female in the low-hormone phase would have to improve her percentile rank by 42% to be as likely to enter a tournament as a female in the high-hormone phase.

The last two columns of Table 9 show that the results concerning females in the low hormonal phase persist when using standard errors that are clustered on the specific experiment session or on the individual for the no information treatment. In this table the dependent variable is the same ordered variable of competitiveness that was used previously. The results are entirely consistent

with our previous findings. Although gender may have a prominent role in explaining females' reluctance to compete in mixed gender settings, it seems that hormonal phase may be a driving factor that needs to be considered in low information settings.

Interestingly, relative performance feedback makes these hormone specific effects disappear. Table 10 provides the results from ordered probit estimations for Treatment 5, where subjects were provided with relative performance information from Treatment 1 prior to making their competitive environment selections. Table 10 shows that when participants are informed of their relative performance compared to other potential competitors, then there is little difference in selection between genders or across the menstrual cycle.

As with the gender differences, we find that after participants are informed of the quality of potential competitors, choice differences across the menstrual cycle become insignificant. We find that choices after feedback mainly depend on the relative performance information provided prior to making the decision and, to a certain extent, on an individual's improvement from Treatment 1 to Treatment 2. Though females' choices to avoid competitive environments are most frequent in the low-hormone phase, these results suggest that this effect seems to be linked with the information available about the quality of potential competitors.

As discussed previously, there is a cost associated with high ability individuals avoiding competitive settings, and with low ability individuals choosing tournaments. We find that females in the low-hormone phase make more costly mistakes than both high phase females and males. The average expected value losses for males, low phase and high phase females in Treatment 4 are shown in Table 11. Low phase females sacrifice the greatest amount of expected value from making suboptimal choices, \$8.31. The expected value losses for high phase females and males are \$6.42 and \$4.89. The differences between low and high phase females and between males and high phase females are not statistically significant, but low phase females make more costly choices than males at the 5% significance level.

These results show that menstrual phase and the corresponding hormone levels are correlated with competitive choices, but only if the strength of the competition is not known. If there is little information about the relative abilities, then females in the low-hormone phase make more costly decisions than males and non-low phase females. But there are no significant differences in expected value losses between genders or between different hormonal phases for females if good relative performance information is available.

Table 10: Ordered Probit: Hormone Effects After Feedback (Treatment 5)

Sample	(1) All	(2) Females	(3) Risk
Female	-0.13 (0.20)		0.10 (0.24)
Low Phase	0.13 (0.25)	0.12 (0.27)	-0.12 (0.29)
Confidence (T1)	0.32 (0.31)	0.21 (0.53)	0.54 (0.36)
Improve (T2)	0.92*** (0.26)	1.06*** (0.40)	0.49 (0.33)
%-tile Rank (T1)	2.61*** (0.35)	2.63*** (0.55)	2.33*** (0.38)
Risk Controls	No	No	Yes
Demographic Controls	Yes	Yes	Yes
Observations	328	155	211
ll	-303.9	-143.3	-183.4
chi2	93.82	45.89	75.79

Note: 45 low phase females could be identified for data analysis.

Note: Estimations include individual random effects.

Standard errors in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table 11: Average Loss (\$) from Suboptimal Choices in Treatment 4

	Mean	Std. Error
Male	4.89	0.70
Female Non-Low	6.42	1.26
Female Low	8.31	2.52

### 3 Discussion

We show that the gender differences in competitive choices in mixed gender groups that have been reported in other studies are robust to a variety of protocol changes, including different tasks and variations in the degree of competitiveness of the available choices. We then show that female choices to compete vary across the menstrual cycle, and that in the low-hormone phase females are less likely to enter tournaments than during a high-hormone phase. Females in the high-hormone phase are almost as likely to compete as are men. We find that the gender difference the female choice differences across the menstrual cycle are removed with relative performance feedback. We show the feedback drives high ability females towards more competitive environments, and low ability males towards less competitive choices.

Research on the gender gap in labor markets has focused on both the wage gap and the gap between females and males entering certain occupations. In our experiment, when we consider the cost of suboptimal choices made by both genders in treatments without relative performance feedback, we find that suboptimal decisions by females are 35% more costly than those of males, a result primarily driven by a 14% lower probability of females entering competitions. The cost differences are higher than the choice difference because some of the most able women are choosing not to enter these competitions, even though doing so would lead to large economic returns.

We find that females in the low-hormone phase have a 16 percentage point lower probability of competing than females in the high phase, and a 26 percentage point lower probability than males. Again the economic cost of suboptimal choices due to these menstrual cycle effects are substantial, once one considers that many of these are high ability females with high probabilities of winning a competition. Compared to perfect sorting, females in the low-phase lose approximately 70% more than males from making suboptimal choices, while females in the high hormone-phase lose 31% more than males. Thus, the gender gap effect of the hormonal cycle can be quite large relative to the raw choice difference, because it is the most able females that are not entering the competitions. While these are large percentage differences, these results are still small relative to the payoffs from choices found in the regular labor market. Little research has been done to explore menstrual cycle effects in labor markets, where it is often more difficult to gather data on similar choices to those used in experiments along with the hormonal phase. One example can be found in Ichino and Moretti (2009), who use attendance data for a large Italian bank and find that absenteeism among females seems to be linked to the 28-day menstrual cycle.

An illustrative example of the potential size of the information effect can be found in the labor market from a U.S. Department of Labor study Solis and Hall (2012). In 2011, approximately 41% of management positions were filled by females, while the females in these jobs earned approximately 28% less than males. If high ability females are less likely to enter these competitive positions than high ability males, then this choice difference provides a potential explanation for the wage gap. One could argue that in the field of management it may be difficult to quantify performance and compare different individuals, as performance evaluations might be based on subjective appraisals. Therefore, we also look at a job where quality might be more objectively measured, i.e. computer programming and software engineering. Females make up approximately 22% of the computer programming workforce and 22% of the software engineer workforce, while the gender wage gap

in those positions is only 5% and 9% (respectively) lower than males. A possible explanation for this could be that high ability females do enter these labor markets, either because the positions are less like competitions or because feedback is more available. So even though there is a large difference in the gender composition of this workforce the wage gap remains low. Obviously our work does not prove this is the reason for these differences in wage gaps, but the differences are consistent with what we find in terms of selection effects driving large earnings differences.

The fact that information removes any significant differences in competitive choices has important economic implications. First, this result suggests that decisions by females not to compete may be more affected by the informational environment and the ambiguity of relative abilities than by any persistent preference for or against competition. Second, since females in the low-hormone phase are less willing to compete than females in the high-hormone phase, this suggests that the hormones linked with the phases of the menstrual cycle may be affecting areas of the brain that help evaluate the rewards and risks associated with these competitive tasks, as suggested by the findings of Dreher et al. (2007) and Vliet (2001). Elevated progesterone could therefore reduce the disutility of worrying about the outcome of a competition – particularly before feedback information. Estradiol has been found to increase the density of serotonin binding sites in areas of the brain that are linked with evaluating rewards (Fink et al. 1996). Females may experience greater sensitivity to potential rewards during elevated levels of estrogen, which could lead them to be more competitive Stanton and Schultheiss (2007). Thus, our results on economic choices are consistent with previous endocrinological studies on hormones and behavior.

Our finding that women’s competitive choices increase during a part of the menstrual phase when progesterone and estrogen are high is noticeably different than the result reported in the only other economics paper to look at this issue. Buser (2012) finds that when females compete against other females, females in menstrual phases with high levels of progesterone are less competitive. One major difference between our studies is that we use mixed-gender groups. Studies have found that female competitive behavior in single gender groups is very different from that in mixed gender groups, with females being much more willing to compete when the group is all female (Gupta et al. 2005, Grosse and Reiner 2010). Being in a same gender group may promote the effect of progesterone on group affiliation motives reported in Schultheiss et al. (2004). As the gender composition experiments show, competitive decisions are context dependent, and the neuroendocrinological literature van Anders and Watson (2006) emphasizes that the behavioral ef-

ffects of hormones are also context dependent. Thus, the gender composition of the particular labor market being examined may interact with the effects of hormones. Another difference is that we use a within-subjects design, which is particularly important in examining the effects of hormonal variations that are inherently within-subjects. A further difference between these studies is that our study includes a larger number of female subjects in both high and low-hormone phases. One potential disadvantage of the within-subjects design is that experiences in the first session might affect choices in the second. We use two different tasks, with the session order randomized to minimize this problem, and we find the results are robust to order and tasks.

The existence of regular and predictable within-subjects variation in behavior across the menstrual cycle suggests that there is a biological component to the gender difference in competitive choices. We cannot conclude that hormones explain the male female differences for competitive choices entirely— there are too many other things that vary across genders, including the different responses to feedback we have demonstrated. Instead, we show that these predictable biological variations for females are linked with variations in behavior across time. But these variations in choices can be reduced or removed through the use of hormonal contraceptives and through performance feedback.

Interestingly, hormonal contraceptives allow for control of the timing of the low-hormone phase, and in fact, there is no medical reason for the low hormone (placebo) phase of the hormonal contraceptive regimen Anderson and Hait (2003). As a leading textbook on clinical gynecological endocrinology states: *Monthly bleeding, periodic bleeding, or no bleeding – this is an individual woman’s choice (Speroff and Fritz, 2005, pg 908)*. It has been suggested that the placebo and withdrawal bleeding was a marketing effort to make the birth control pill seem less novel and more acceptable Coutinho and Segal (1999). A number of studies have found that the majority of females would prefer not to experience the monthly bleeding caused by low hormone levels or the naturally occurring menstrual cycle (Andrist et al. 2004, Makuch et al. 2012, Lakehomer et al.2012), and some new contraceptives eliminate the placebo pills and ensure that the low-hormone phase does not occur Anderson and Hait (2003); Archer (2006). A study of Australian female university students found that 66% of those on oral contraceptives had engaged in prolonged pill-use at some point in their lives to avoid monthly bleeding; the majority of such avoidance was done for reasons of convenience (Greig et al. 2010). Approximately, half of oral contraceptive users among deployed

female aviation personnel and females deployed for combat operations took the pill continuously (Powell-Dunford et al. 2009, Powell-Dunford et al. 2011).

The combination of easily available and commonly used methods for manipulating hormone levels and the menstrual cycle with the evidence we provide on behavioral differences across the cycle raises interesting economic questions. Women who know that they are subject to the behavioral variations we have found might take steps to increase or decrease their competitive choices by changing when they make decisions involving competition, such as job choices and college applications.

Since information about relative ability removes all behavioral differences across the menstrual cycle and between genders, it would be natural to suggest that the differences in competitive choices might be the result of differences in overconfidence. But we find gender differences in competitive choices in both math and word tasks. In the word tasks, confidence levels of females and males are the same and a gender difference in choices persists. In the math task the gender difference persists even after controlling for confidence. The information about relative abilities may help subjects form more accurate beliefs about the possibilities of succeeding in competitive situations, but since females hold similar overconfident beliefs as males in word tasks, we would expect females to choose in a similar fashion as males. Instead, females choose not to compete in low information settings, irrespective of confidence or the type of task.

Another possible explanation for the differences in choices across gender with and without feedback would be a gender difference in risk aversion. However, when we control for risk aversion the gender difference in choices still remains in the uninformed treatment. Furthermore we find risk aversion is not significant in explaining competitive choices in the full information treatment. Another explanation that would be consistent with our results would be a gender difference in ambiguity aversion. Little research has been done on gender differences in ambiguity aversion, and the results are mixed.

Firms, governments, and schools sometimes implement affirmative action policies to encourage females to apply for competitive jobs and scholarships. Affirmative action policies typically focus efforts on recruiting females or changing the acceptance or promotion process to favor females. Niederle et al. (2009) report on experiments showing that affirmative action can encourage competition by females with low efficiency costs. The performance feedback result from our study suggests a simple alternative to affirmative action— providing better information about relative abilities. We

find that information on relative ability plays a large role in choices to compete. Information reduces gender differences, reduces differences across the menstrual cycle, and improves sorting by ability. Schools and many large firms already collect relative performance rankings about students or employees Pfeffer and Sutton (2006), and our work suggests that making these rankings more available to these agents and emphasizing relative comparisons will promote an environment where the best workers, both females and males, will seek out more competitive positions. These experimental results are corroborated with empirical evidence showing that males and females respond differently to feedback in professional tennis tournaments Wozniak (2012). A caveat to the experimental evidence is that we find that when we repeat the choice experiment a few weeks after providing feedback (albeit for a different task) the gender difference returns when no relative performance feedback is provided to subjects about the task. This suggests that feedback may need to be sustained to be effective.

Our results support the idea that biological factors play an important role in explaining the gender difference for competitions. We find a large gender difference, and a large difference within females that is correlated with hormonal variations. But these differences between hormonal phases and between genders no longer become significant with detailed performance feedback. Although biology plays a role in explaining gender differences for competition, the effects of information are much greater and lead individuals to make more efficient choices regardless of biological differences.

## References

- Aden, Ulrike, Claudia Jung-Hoffmann, and Herbert Kuhl. 1998. "A randomized cross-over study on various hormonal parameters of two triphasic oral contraceptives." *Contraception* 58 (2):75–81.
- Anderson, F.D. and Howard Hait. 2003. "A multicenter, randomized study of an extended cycle oral contraceptive." *Contraception* 68 (2):89–96.
- Andrist, Linda C., Raquel D. Arias, Deborah Nucatola, Andrew M. Kaunitz, B. Lynn Musselman, Suzanne Reiter, Jennifer Boulanger, Linda Dominguez, and Steven Emmert. 2004. "Women's and providers' attitudes toward menstrual suppression with extended use of oral contraceptives." *Contraception* 70 (5):359–363.
- Apicella, Coren L., Anna Dreber, Benjamin C. Campbell, Peter B. Gray, Moshe Hoffman, and Anthony C. Little. 2008. "Testosterone and financial risk preferences." *Evolution and Human Behavior* 29 (6):384–390.
- Archer, David F. 2006. "Menstrual-cycle-related symptoms: a review of the rationale for continuous use of oral contraceptives." *Contraception* 74 (5):359–366.
- Arellano, Manuel and Bo Honoré. 2001. "Panel data models: some recent developments." *Handbook of Econometrics* 5:3229–3296.
- Association, American College Health American College Health. 2009. "ACHA National College Health Assessment II: Reference Group Data Report." Tech. rep., ACHA.
- Bertrand, Marianne and Kevin F. Hallock. 2001. "The gender gap in top corporate jobs." *Industrial and Labor Relations Review* :3–21.
- Bethea, Cynthia L., Nick Z. Lu, Chrisana Gundlah, and John M. Streicher. 2002. "Diverse Actions of Ovarian Steroids in the Serotonin Neural System." *Frontiers in Neuroendocrinology* 23 (1):41–100.
- Borghans, Lex, Bart H.H. Golsteyn, James J. Heckman, and Huub Meijers. 2009. "Gender Differences in Risk Aversion and Ambiguity Aversion." *Journal of the European Economic Association* 7 (2-3):649–658.
- Brown, Charles. 1990. "Firms' choice of method of pay." *Industrial and Labor Relations Review* :165–182.
- Burnham, T.C. 2007. "High-testosterone men reject low ultimatum game offers." *Proceedings of the Royal Society B: Biological Sciences* 274 (1623):2327.
- Buser, Thomas. 2012. "The impact of the menstrual cycle and hormonal contraceptives on competitiveness." *Journal of Economic Behavior & Organization* 83 (1):1–10.
- Chakraborty, Tandra R., Laurie Ng, and Andrea C. Gore. 2003. "Age-Related Changes in Estrogen Receptor {beta} in Rat Hypothalamus: A Quantitative Analysis." *Endocrinology* 144 (9):4164.
- Chen, Yan, Peter Katuscak, and Emre Ozdenoren. 2009. "Why Can't a Woman Bid More Like a Man?" *Working Paper, Ann Arbor* 1001:48109–1220.
- Coates, JM and J. Herbert. 2008. "Endogenous steroids and financial risk taking on a London trading floor." *Proceedings of the National Academy of Sciences* 105 (16):6167.

- Coutinho, Elsimar M. and Sheldon J. Segal. 1999. *Is menstruation obsolete?* Oxford University Press, USA.
- Creinin, Mitchell D., Sharon Keverline, and Leslie A. Meyn. 2004. “How regular is regular? An analysis of menstrual cycle regularity.” *Contraception* 70 (4):289–292.
- Datta Gupta, Nabanita, Anders Poulsen, and Marie-Claire Villeval. 2005a. “Do (wo) men prefer (non-) competitive jobs.” Tech. rep., GATE Working Paper.
- . 2005b. “Male and Female Competitive Behavior-Experimental Evidence.” *IZA Working Paper* .
- Dohmen, Thomas and Armin Falk. 2011. “Performance pay and multidimensional sorting: Productivity, preferences, and gender.” *The American Economic Review* 101 (2):556–590.
- Dreher, Jean-Claude, Peter J. Schmidt, Philip Kohn, Daniella Furman, David Rubinow, and Karen Faith Berman. 2007. “Menstrual cycle phase modulates reward-related neural function in women.” *Proceedings of the National Academy of Sciences* 104 (7):2465.
- Eckel, Catherine C. and Phillip J. Grossman. 2008. “Men, women and risk aversion: Experimental evidence.” *Handbook of experimental economics results* 1:1061–1073.
- Fernandez, Guillén, Susanne Weis, Birgit Stoffel-Wagner, Indira Tendolkar, Markus Reuber, Stefan Beyenburg, Peter Klaver, Jurgen Fell, Armin de Greiff, Jurgen Ruhlmann et al. 2003. “Menstrual cycle-dependent neural plasticity in the adult human brain is hormone, task, and region specific.” *Journal of Neuroscience* 23 (9):3790.
- Fink, George, Barbara E. Sumner, Roberta Rosie, Oliver Grace, and John P. Quinn. 1996. “Estrogen control of central neurotransmission: effect on mood, mental state, and memory.” *Cell Mol Neurobiol* 16 (3):325–44.
- Gagnidze, Khatuna and Donald W. Pfaff. 2009. “Sex on the Brain.” *Cell* 139 (1):19–21.
- Gneezy, Uri, Kenneth L. Leonard, and John A. List. 2009. “Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society.” *Econometrica* 77 (5):1637–1664.
- Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. “Performance in Competitive Environments: Gender Differences.” *Quarterly Journal of Economics* 118 (3):1049–1074.
- Gneezy, Uri and Aldo Rustichini. 2004. “Gender and Competition at a Young Age.” *American Economic Review* 94 (2):377–381.
- Greene, William. 2004. “The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects.” *Econometrics Journal* 7 (1):98–119.
- Greig, Alecia J., Michelle A. Palmer, and Lynne M. Chepulis. 2010. “Hormonal contraceptive practices in young Australian women (25 years) and their possible impact on menstrual frequency and iron requirements.” *Sexual & Reproductive Healthcare* 1 (3):99–103.
- Grosse, Niels D. and Gerhard Riener. 2010. “Explaining Gender Differences in Competitiveness: Gender-Task Stereotypes.” *Jena Economic Research Papers* 2010:017.
- Gunther, C., N.A. Ekinçi, C. Schwierien, and M. Strobel. 2008. “Women can’t jump?—An experiment on competitive attitudes and stereotype threat.” Tech. rep., Mimeo.

- Günther, Christina, Neslihan A. Ekinçi, Christiane Schwieren, and Martin Strobel. 2010. “Women cant jump? An experiment on competitive attitudes and stereotype threat.” *Journal of Economic Behavior & Organization* 75 (3):395–401.
- Holt, Charles A. and Susan K. Laury. 2002. “Risk Aversion and Incentive Effects.” *American Economic Review* 92 (5):1644–1655.
- Ichino, Andrea and Enrico Moretti. 2009. “Biological Gender Differences, Absenteeism and the Earnings Gap.” *American Economic Journal: Applied Economics* 1 (1):183–218.
- Jirjahn, Uwe and Gesine Stephan. 2004. “Gender, piece rates and wages: evidence from matched employer-employee data.” *Cambridge Journal of Economics* 28 (5):683–704.
- Lakehomer, Hannah, Paul Kaplan, David Wozniak, and Chris Minson. 2012. “Characteristics of Menstrual Cycle Manipulation With Combined Hormonal Contraception in a University Student Population.” *Fertility and Sterility* 97 (3):S30.
- Makuch, Maria Y., Maria J. Duarte-Osis, Karla S. de Pádua, Carlos Petta, and Luis Bahamondes. 2012. “Opinion and experience of Brazilian women regarding menstrual bleeding and use of combined oral contraceptives.” *International Journal of Gynecology & Obstetrics* .
- Malison, Robert T., Lawrence H. Price, Robert Berman, Christopher H. van Dyck, Gregory H. Pelton, Linda Carpenter, Gerard Sanacora, Michael J. Owens, Charles B. Nemeroff et al. 1998. “Reduced brain serotonin transporter availability in major depression as measured by [123I]-2 $\beta$ -carbomethoxy-3 $\beta$ -(4-iodophenyl) tropane and single photon emission computed tomography.” *Biological Psychiatry* 44 (11):1090–1098.
- Mayr, Ulrich, David Wozniak, Casey Davidson, David Kuhns, and William T. Harbaugh. 2012. “Competitiveness across the life span: The feisty fifties.” *Psychology and aging* 27 (2):278.
- McEwen, Bruce. 2002. “Estrogen Actions Throughout the Brain.” *Recent Progress in Hormone Research* 57 (1):357–384.
- Moore, Evan and Catherine Eckel. 2003. “Measuring Ambiguity Aversion. Virginia Tech.”
- Mosher, William D., Gladys M. Martinez, Anjani Chandra, Joyce C. Abma, and Stephanie J. Willson. 2004. “Use of contraception and use of family planning services in the United States: 1982–2002.” *Advance data from vital and health statistics* 350:1.
- Niederle, Muriel, Carmit Segal, and Lise Vesterlund. 2009. “How costly is diversity? Affirmative action in light of gender differences in competitiveness.” *NBER Working Paper* .
- Niederle, Muriel and Lise Vesterlund. 2007. “Do Women Shy Away from Competition? Do Men Compete Too Much?” *Quarterly Journal of Economics* 122 (3):1067–1101.
- . 2010. “Explaining the Gender Gap in Math Test Scores: The Role of Competition.” *The Journal of Economic Perspectives* 24 (2):129–144.
- Pearson, Matthew and Burkhard C. Schipper. 2009. “Menstrual cycle and competitive bidding.” *MPRA Paper* .
- Pfeffer, Jeffrey and Robert I. Sutton. 2006. “Evidence-based management.” *Harvard Business Review* 84 (1):62.

- Platt, Michael L. and Scott A. Huettel. 2008. "Risky business: the neuroeconomics of decision making under uncertainty." *Nature Neuroscience* 11 (4):398–403.
- Polachek, Solomon William. 1981. "Occupational self-selection: A human capital approach to sex differences in occupational structure." *The Review of Economics and Statistics* :60–69.
- Powell-Dunford, Nicole, Amanda S. Cuda, Jeffrey L. Moore, Mark S. Crago, and Patricia A. Deuster. 2009. "Menstrual suppression using oral contraceptives: survey of deployed female aviation personnel." *Aviation, space, and environmental medicine* 80 (11):971–975.
- Powell-Dunford, Nicole C., Amanda S. Cuda, Jeffrey L. Moore, Mark S. Crago, Amanda M. Kelly, and Patricia A. Deuster. 2011. "Menstrual suppression for combat operations: Advantages of oral contraceptive pills." *Women's Health Issues* 21 (1):86–91.
- Schultheiss, Oliver C., Michelle M. Wirth, and Steven J. Stanton. 2004. "Effects of affiliation and power motivation arousal on salivary progesterone and testosterone." *Hormones and Behavior* 46 (5):592–599.
- Shepard, Kathryn N., Vasiliki Michopoulos, Donna J. Toufexis, and Mark E. Wilson. 2009. "Genetic, epigenetic and environmental impact on sex differences in social behavior." *Physiology & behavior* 97 (2):157–170.
- Sinha-Hikim, Indrani, Stefan Arver, Gildon Beall, Ruoqing Shen, Mario Guerrero, Fred Sattler, Cecilia Shikuma et al. 1998. "The use of a sensitive equilibrium dialysis method for the measurement of free testosterone levels in healthy, cycling women and in human immunodeficiency virus-infected women." *Journal of Clinical Endocrinology & Metabolism* 83 (4):1312.
- Solis, Hilda and Keith Hall. 2012. "Women in the Labor Force: A Databook." Tech. Rep. 1034, U.S. Department of Labor, U.S. Bureau of Labor Statistics.
- Speroff, Leon and Marc A. Fritz. 2005. *Clinical Gynecologic Endocrinology and Infertility*. Lippincott Williams & Wilkins.
- Stanton, Steven J. and Oliver C. Schultheiss. 2007. "Basal and dynamic relationships between implicit power motivation and estradiol in women." *Hormones and Behavior* 52 (5):571–580.
- Stricker, Reto, Raphael Eberhart, Marie-Christine Chevailler, Frank A. Quinn, Paul Bischof, and Rene Stricker. 2006. "Establishment of detailed reference values for luteinizing hormone, follicle stimulating hormone, estradiol, and progesterone during different phases of the menstrual cycle on the Abbott ARCHITECT analyzer." *Clinical Chemistry and Laboratory Medicine* 44 (7):883.
- van Anders, Sari M. and Neil V. Watson. 2006. "Social Neuroendocrinology: Effects of Social Contexts and Behaviors on Sex Steroids in Humans." *Human Nature* 17 (2):212–237.
- Vliet, Elizabeth L. 2001. *Screaming to be heard: hormone connections women suspect and doctors still ignore*. New York, NY: M. Evans and Co.
- Wozniak, David. 2012. "Gender Differences in a Market with Relative Performance Feedback: Professional Tennis Players." *Journal of Economic Behavior & Organization* 83 (1):158–171.
- Wu, Melody V., Devanand S. Manoli, Eleanor J. Fraser, Jennifer K. Coats, Jessica Tollkuhn, Shin-Ichiro Honda, Nobuhiro Harada, and Nirao M. Shah. 2009. "Estrogen masculinizes neural pathways and sex-specific behaviors." *Cell* 139 (1):61–72.

Zethraeus, Niklas, Ljiljana Kocoska-Maras, Tore Ellingsen, Bo von Schoultz, Angelica L. Hirschberg, and Magnus Johannesson. 2009. "A randomized trial of the effect of estrogen and testosterone on economic behavior." *Proceedings of the National Academy of Sciences* 106 (16):6535.