ESSAYS IN EXPERIMENTAL ECONOMICS: EXAMINING THE EFFECTS OF AMBIGUITY AND COMPETITION

by

DAVID WOZNIAK

A DISSERTATION

Presented to the Department of Economics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2010

University of Oregon Graduate School

Confirmation of Approval and Acceptance of Dissertation prepared by:

David Wozniak

Title:

"Essays in Experimental Economics: Examining the Effects of Ambiguity and Competition"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Economics by:

William Harbaugh, Chairperson, Economics
Trudy Cameron, Member, Economics
Van Kolpin, Member, Economics
Christopher Minson, Outside Member, Human Physiology

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

June 14, 2010

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.

An Abstract of the Dissertation of

David Wozniak for the degree of Doctor of Philosophy

in the Department of Economics to be taken June 2010

Title: ESSAYS IN EXPERIMENTAL ECONOMICS: EXAMINING THE EFFECTS OF AMBIGUITY AND COMPETITION

Approved: ______________________________
Dr. William T. Harbaugh

Individuals compete against each other in a variety of different settings. In labor markets they compete for promotion; in athletic tournaments they compete for fixed prizes. Important aspects of competitive choices include the probability of success, expected payoffs, the level of ambiguity regarding success, and preferences to compete. I explore the effects of biology and relative performance feedback in regard to these components in three essays.

In the first essay I use a unique experiment design to measure ambiguity aversion, which can be modified to also control for risk aversion. A measure of ambiguity aversion has value as individuals in labor markets have ambiguous signals about their probabilities of success in competition. Consequently this measure may be used in future experiment designs to control for heterogeneous preferences for ambiguity and to test whether ambiguity affects behaviors differently than risk.

Economic experiments 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. In the second essay I replicate these findings and then show that giving relative performance feedback moves high ability women towards more competitive compensation schemes, moves low ability men towards less competitive compensation schemes, and removes the gender difference in compensation choices. I then examine differences in choices for women, across the menstrual cycle. I find that women in the low-hormone phase of their cycle are less likely to enter tournaments than women in the high-hormone phase. Men are more likely to choose tournaments than women at either stage. There are no significant selection differences between any of these groups after they receive relative performance feedback.

Athletic labor markets provide a unique environment where individuals choose to compete when they have high quality information about their potential competitors. Gender differences for competition have been found to be removed when information about relative abilities is available. In the third essay, to explore the effect of information in a labor market setting, I use a unique data set of approximately 6,000 female and male competitive tennis players during the 2009 season. I focus on whether males and females choose to enter competitive tournaments differently in response to past performance. I find that males continue to compete after performing well in the previous week while females are less likely to compete if they do well. These contrasting behaviors suggest that males and females respond differently to performance feedback.

CURRICULUM VITAE

NAME OF AUTHOR: David Wozniak

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
Dalhousie University, Halifax, Nova Scotia
University of Calgary, Calgary, Alberta

DEGREES AWARDED:

Doctor in Philosophy, Economics, University of Oregon (2010)
Master of Arts in Economics, Dalhousie University (2006)
Bachelor of Arts Honours in Economics, University of Calgary (2002)
Bachelor of Commerce, University of Calgary (2002)

AREAS OF SPECIAL INTEREST:

Experimental Economics
Behavioral Economics
Labor Economics

ACKNOWLEDGEMENTS

I wish to thank my adviser Dr. Bill Harbaugh and the rest of my committee. Dr. Bill Harbaugh has provided me with every opportunity possible and I am grateful to have been lucky enough to work under his direction. His time and patience in dealing with me are unmeasurable and I would not be able to accomplish very much without his help. Dr. Trudy Cameron has been instrumental in steering me through so many of the obstacles that I encountered. Dr. Van Kolpin has allowed me to better understand the problems encountered when dealing with ambiguity. Dr. Chris Minson has played a vital role in helping me understand many of the biological effects that occur in the human body. His support for a topic well off the beaten path is very much appreciated. I would also like to thank Ulrich Mayr, Sarah Wunderlich, and Jack Wozniak; without their efforts this entire research agenda may not have come to fruition.

TABLE OF CONTENTS

Chapter | Page

LIST OF FIGURES

Figure | Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

In the proceeding chapters I explore the effects of information for individual decision making. I first consider how a lack of information for uncertain outcomes may affect individuals for simulated investment decisions. I find a possible experimental design that may help explore ambiguity effects in the future.

In Chapter III, I examine how information about individuals' abilities compared to potential competitors changes how they enter competitive environments. Specifically, I find that such information may remove gender differences for competition. Thus, such relative performance feedback may help eliminate some of the gender differences we find today in labor markets.

In Chapter III I also consider a biological reason for gender differences for competition. Females experience large and fairly predictable hormonal fluctuations across the menstrual cycle or from hormonal birth control use. I show that entry into competitive environments is correlated with these hormonal fluctuations. Thus, predictable biological functions may play a role in explaining some of the heterogeneity or inconsistencies that is observed by individuals in economic markets.

In the following chapter, I examine a market where both men and women compete: the market of professional tennis players. This market consists of very competitive individuals where both men and women compete under similar institutions. Furthermore, these market institutions provide very good relative performance feedback. In an economic laboratory it was found that males and females compete similarly in environments with good relative performance information; therefore, I examine whether men and women choose to enter tournaments in a similar fashion in an actual labor market. I find that even in these settings men and women behave differently. Specifically, the effects of past performance seem to have very different effects for men and women. Consequently, further understanding of the economic consequences of heterogeneous effects to feedback must be understood before using performance feedback as a policy mechanism for both genders to compete efficiently.

CHAPTER II

WTP TO REMOVE AMBIGUITY EFFECTS

2.1 Introduction

In discussions of ambiguity, uncertainty and risk, the terms often get interchanged. For the purposes of this exposition, I will attempt to maintain a certain degree of consistency in how the terms are used. Risk involves using knowable and well defined probabilities to describe a random event. Uncertainty and risk will be used to describe an asset that has an underlying probability distribution containing calculable risks. Thus, uncertainty will refer to a situation where outcomes become unknown due to first or second order probability distributions. Such outcomes can be modeled as lotteries or compound lotteries.

To describe ambiguity, I will use the definition found in Kagel and Roth (1995): ambiguity involves *known-to-be-missing information.* For example, ambiguity may refer to the case where the individual does not have a clear signal as to the second order probability distribution that is compounding an underlying risky asset. Ambiguity can occur if an individual is not sure about what distribution (normal, log-normal, uniform) is being used for a set of known lotteries; such an individual has trouble forming beliefs about the possible distributions regarding an underlying asset from ambiguous signals or past experience.

For example, individuals often encounter some degree of ambiguity about investment assets or potential workers. This ambiguity can occur in the form of a reference letter that ambiguously states that a worker's performance was often discussed by supervisors without stating if the quality of performance was high or low. This ambiguous signal leaves a potential employer wrestling with ambiguity regarding the quality of the potential hire. Another form of ambiguity that may occur is in bluffing situations, such as the game of poker: poker players will bet on an asset (their hand) so as to create an ambiguous signal about the distribution of dealt cards. Once the hands are dealt, it is difficult to read which cards are being played by opponents, even more so once ambiguous signals are sent through betting behavior. A *Texas Hold'em* poker player understands that a deck of cards has a known distribution across 52 cards, but the particular distribution of cards dealt to opponents takes on a different strategic meaning. Players form assumptions about the dealt distribution based on private information (one's own hand of cards) and public signalling in the form of betting. Ambiguous signals may leave the card player with no better assumption than to play the naive distribution of the other 50 cards that was known to exist prior to the hands being dealt.[1]

To measure ambiguity aversion, I use a unique experimental framework to re-examine a classic *urn* problem that has been widely used to characterize ambiguity and uncertainty, the so-called Ellsberg urn (F.H. Knight 1921, D. Ellsberg 1961, L.J. Savage 1972). Previous experiments have examined notions of ambiguity and uncertainty by observing subjects' revealed preferences concerning bets from two possible urns or in recording subjects' values of lotteries between an uncertain urn

[1]In *Texas Hold'em* each player receives two cards each.

and an ambiguous urn (C.R. Fox & A. Tversky 1995, C. Heath & A. Tversky 1991). Often these urns consist of something similar to the following:

1. An urn with a known mixture of 10 red and black balls (i.e. 5 red and 5 black). (Uncertain Urn)

2. An urn with an unknown mixture of 10 balls for which any could be red or black. (Ambiguous Urn)

Subjects must bet on pulling out a ball of a certain color from one of these urns and they have the option of picking whether that ball will be pulled out of the Uncertain or Ambiguous urn.

Results show that most individuals prefer to bet on the Risky urn instead of the Ambiguous urn. In a recent study, Halvey (2007) uses the Becker-DeGroot-Marschak (1964) mechanism to examine which theory best describes the valuations of these urns by individuals. Halevy (2007) focuses on the relationships between individual attitudes towards ambiguity and compound objective lotteries. Previously, Lichtenstein et al. (1978) found that individuals typically have biases that lead them to overestimate the frequencies of low probability events and underestimate high probability events. These types of biases were also found to exist in exploring the accuracy of mortality risk among a general population (JK Hakes & WK Viscusi 2004). These types of biases may impact how individuals evaluate ambiguous situations because if the risks are well defined, and agents are informed of them, then such biases will be removed.

This experiment focuses on individuals' willingness to pay (WTP) to remove ambiguity. As opposed to providing subjects with different urns and lotteries, I take a different approach to examine the effects of ambiguity. I use an experiment where subjects must invest in an ambiguous asset, but the novelty of this experimental design is that subjects can spend part of their possible investment to remove some or all of the effects of ambiguity associated with the asset. To my knowledge, there have been only two experiments that use sampling of an urn by individuals to remove the effects of ambiguity (J.S. Chipman 1960, G. Gigliotti & B. Sopher 1996).[2] In these experiments, the sampling was costless and the effect of sampling was the topic of study. My WTP task, in combination with the Holt and Laury (2002) risk aversion task, allows me to examine whether risk aversion is correlated with aversion to ambiguity. Furthermore, the design provides the tools necessary to calculate lower bounds on the ex-ante beliefs that individuals form as a consequence of sampling the urn.

In contrast to a recent study by Moore and Eckel (2003), who measure ambiguity aversion as an aversion to second-order probability distributions, I attempt to create a more ambiguous situation. One that leaves subjects with missing information about the second-order probability distribution used to generate a risky asset. Eckel and Moore use a range of probabilities and dollar amounts for a risky asset and allow subjects to choose between the fuzzy probability or fuzzy dollar amount and a "sure thing" dollar amount. This approach provides a measure of attitudes towards compound probabilities and fuzzy dollar amounts and not necessarily towards ambiguity.

---

[2]Chipman only used 10 subjects

I focus on how individuals pay to remove some or all or none of the effects of ambiguity involved with the asset. In a computerized experiment individuals are given an urn with 10 balls that could be black and red, where the mixture of black and red balls is unknown. Individuals are not informed as to the random generating process used to create the mix of black and red balls in the urn, thus creating an ambiguous situation regarding the distribution that compounds a gamble. A priori, if one has symmetric beliefs, then a 0.5 probability should be assigned to drawing a black or red ball. Individuals are given a budget (w) for the investment task from which they can pay for a costly sample (c per ball) from the urn. An individual can sample the urn before betting on whether they will remove a red or black ball. For example, if an individual samples 3 balls then 3 balls are removed from the urn simultaneously and shown to the subject for a cost of $3c$. The balls are then placed back in the urn and the subject must bet on a color with the full remaining amount of his investment budget (from the example the amount remaining is $w - 3c$). If an individual correctly guesses the color of the ball that gets drawn, after being given the opportunity to pay to sample the urn, then he receives twice the amount of the budget that is left after he pays to sample the urn's contents.

During experiment sessions, individuals also participate in a task used to determine their risk preferences. The Holt and Laury (2002) method is used to estimate coefficients of relative risk aversion (CRRA). I find that a large number of individuals are willing to pay to remove some of the effects of ambiguity. Only 19% of subjects were willing to gamble on the urn without paying for a costly sample beforehand. I find that the estimated risk aversion coefficients play a significant role in how individuals' are willing to pay to remove the effects of ambiguity. Assuming that prior beliefs are heterogeneous across individuals,[3] it seems that risk averse individuals over-weight the tails of probability distributions of outcomes; in other words, they seem to believe very unlikely outcomes are likely to occur.

The following section will provide a description of subjects, this will be followed by a detailed explanation of the risk aversion experiment. The third section will focus on the experiment concerning the WTP to remove ambiguity. The fourth section will provide a discussion of results and will review the apparent beliefs of individuals.

2.2 Subjects

Experiment participants consisted of seventy undergraduate students (11 females and 59 males) who were enrolled in an intermediate-level economics class at a mid-size public university.[4] Subjects' ages were between 19 and 34 with a mean of 22 and a standard deviation of 2.3 years. There were three experiment sessions that lasted approximately thirty minutes each, and the numbers of students in each session were: 26, 27, and 17 respectively.[5]

Subjects were paid $2.35 (SD: 1.36) on average for their task performance with a minimum payout of $0.10 and a maximum of $4 per session. The total amount paid out for these experiment sessions was $164.70. Sessions consisted of 3 different types of experiments and subjects were made aware that their payoffs would depend on

---

[3]These prior beliefs come from a subjective expected utility model (Savage 1972)

[4]73 students participated, but three were unable to complete the entire session or had erroneous self-recording of some of their key data.

[5]Sessions took place during class time though participation was completely voluntary.

their performance in those experiments as well as some random luck. Subjects were informed that their payoffs would be based on just one of the experiments, which was randomly chosen.

2.3 Risk Aversion Experiment

The first experiment in which subjects participated involved a simple task designed to extract a measure of risk aversion for each participant. This method was used by Holt and Laury (2002) to estimate coefficients of relative risk aversion (CRRA). I use the same procedure to estimate a CRRA for each individual in this study. But due to the findings of Holt and Laury concerning different risk averse behaviors for different monetary amounts, I estimate these coefficients using monetary amounts that were similar for each experiment during the session.[6]

Individuals were given a table with a list of ten decisions to make. The decisions involve a choice between a safe investment with payoffs of $2 in the good state and $1.50 in the bad state, which is called Option A. The other option (Option B) is a risky investment with payoffs of $4 in the good state and $0.10 in the bad state. The ten decisions vary in terms of the probabilities of a good state and bad state occurring. For the first decision the probability of the good state is equal to 0.1 and the probability of a bad state is 0.9. For each subsequent decision the probability of the good state increases by 0.1 and the probability of a bad state decreases by 0.1.

[6]In their study, Holt and Laury showed that the a simple CRRA utility function was inadequate. Individuals behaved differently with higher payoffs; this difference was exhibited by an increase in the CRRA with higher payoffs. This result suggested that the CRRA was not constant across payoffs. Instead, they used a hybrid "power-expo" function (A. Saha 1993) to model risk preferences, based on the methods used in this experiment I am unable to do the same.

Thus, rational individuals are expected to choose option A for the first few decisions before switching to option B as the probability of the bad state decreases.[7]

Similar to Holt and Laury (2002), I estimate a risk aversion paremeter using a CRRA utility function for money (x) as follows:

$$u(x) = \frac{x^{1-r}}{1-r} \tag{2.3.1}$$

Assuming the utility function for individuals really takes this functional form, a risk loving individual has $r < 0$, a risk neutral person has $r = 0$, and a risk averse person has $r > 0$. The payoffs for the lottery in the experiment are such that a risk neutral person should choose four safe options (Option A) before switching over to the risky option (Option B) for the fifth decision. Participants were taken through a demonstration of the experiment and were shown how the possible payoffs would be chosen prior to filling in their own risk aversion experiment sheet.

Figure 1 provides the estimates for the maximum CRRA for switching to Option B for a given decision. If the CRRA of an individual is higher for a given decision than the coefficient shown in the far right column of Figure 1 then the individual is more risk averse than the coefficient and will refrain from switching until a later decision.

Table 1 shows the frequency of switching for all experiment participants. In general, subjects seem to be risk averse because approximately 76% of subjects switch

[7]There was a positive probability that subjects were paid for these choices so these were real choices and not just stated choices.

FIGURE 1. Risk Aversion Experiment: Expected Value Differences and CRRA Estimates From Switching

Decision	Option A	Option B	Expected Payoff Diff (EVA-EVB)	CRRA Max. to switch to B
1	$2.00 if the die is 1 $1.50 if the die is 2-10	$4.00 if the die is 1 $0.10 if the die is 2-10	1.06	-1.446
2	$2.00 if the die is 1-2 $1.50 if the die is 3-10	$4.00 if the die is 1-2 $0.10 if the die is 3-10	0.72	-0.761
3	$2.00 if the die is 1-3 $1.50 if the die is 4-10	$4.00 if the die is 1-3 $0.10 if the die is 4-10	0.38	-0.346
4	$2.00 if the die is 1-4 $1.50 if the die is 5-10	$4.00 if the die is 1-4 $0.10 if the die is 5-10	0.04	-0.033
5	$2.00 if the die is 1-5 $1.50 if the die is 6-10	$4.00 if the die is 1-5 $0.10 if the die is 6-10	-0.30	0.233
6	$2.00 if the die is 1-6 $1.50 if the die is 7-10	$4.00 if the die is 1-6 $0.10 if the die is 7-10	-0.64	0.48
7	$2.00 if the die is 1-7 $1.50 if the die is 8-10	$4.00 if the die is 1-7 $0.10 if the die is 8-10	-0.98	0.73
8	$2.00 if the die is 1-8 $1.50 if the die is 9-10	$4.00 if the die is 1-8 $0.10 if the die is 9-10	-1.32	1.012
9	$2.00 if the die is 1-9 $1.50 if the die is 10	$4.00 if the die is 1-9 $0.10 if the die is 10	-1.66	1.397
10	$2.00 if the die is 1-10	$4.00 if the die is 1-10	-2	18.179

to Option B between decisions 5 and 7. This switching behavior suggests that these subjects have relative risk aversion coefficients between 0.0 and 0.7. Approximately 19% of the sample exhibit risk loving behavior and only two subjects wait until decision 10 to switch, suggesting a very extreme form of relative risk aversion.

For the most part subjects seemed to behave rationally. Only 4 subjects reversed their behavior after switching to Option B.[8] These individuals chose Option B for an early decision and then for a later decision they reversed their behavior and switched back to Option A. This behavior cannot be considered very rational, though the possibility exists that in brief time period these individuals' relative risk aversion parameters were fluctuating.[9] At times, the individuals who reversed behavior are

[8]In Holt and Laury (2002) reversals were as high 13% (28 of 212) for some treatments.

[9]Though this behavior is possible it is not very plausible. Instead, I believe subjects did not understand the experiment or were not motivated enough by the economic consequences.

TABLE 1. Frequency of Risk Aversion Coefficients

Switch B	CRRA	Freq.	%	Cum.
2	-0.761	2	2.86	2.86
3	-0.346	7	10.00	12.86
4	-0.033	4	5.71	18.57
5	0.233	21	30.00	48.57
6	0.480	19	27.14	75.71
7	0.730	13	18.57	94.29
8	1.012	2	2.86	97.14
10	18.179	2	2.86	100.00
Total		70	100.00	

excluded from estimations when these reversals are included in any model then an indicator variable helps identify them as being different from the rest of the sample.

TABLE 2. Summary Statistics for Risk Aversion

Stats	Switch B	CRRA	Reversal
Mean	5.53	0.83	0.06
SD	1.57	3.02	0.23
Max	10	18.18	1
Min	2	-0.76	0

Entire sample of 70 subjects.

The summary statistics for the risk aversion experiment are shown in Table 2. Most individuals switch in the vicinity of decisions 5 and 6 and on average this group of subjects has an approximate relative risk aversion parameter of 0.83. The average relative risk aversion parameter is skewed by two individuals who appear to have a risk aversion parameter such that $r \in (1.40, 18.18)$, which is a very extreme form of risk aversion. After removing these individuals from the sample, I find that the average CRRA is 0.315 (SD 0.38). This implies that the mean CRRA is within the range of [0.23,0.48]. The removal of the two extremely risk averse candidates leads to an average CRRA that is more in line with the average switching behavior. This correction is important as using the entire sample for estimation purposes to measure CRRA effects can lead to very implausible results.

TABLE 3. Summary Statistics for Individual Characteristics

Variable	**Mean**	**Std. Dev.**	**N**
Female	0.16	0.37	70
Age	22.15	2.29	69
Econometrics	0.15	0.362	72
Session	1.87	0.78	70

Differences in observations are due to skipped questions by subjects.

Table 3 provides the summary statistics for individual characteristics that may be relevant to risk aversion, which include gender, age, whether they are currently enrolled in an econometrics class, and session that individual self selected to attend.[10]

Subjects could attend one of three possible sessions that were held in consecutive order; subjects were free to self-select into sessions until they became full.[11] The possibility exists that individuals who finished earlier sessions would pass by individuals attending the subsequent session. There may have been communication between individuals from different sessions, which could have affected people in later

[10]One individual chose not to report his or her age. There are two individuals who are included in the summary statistics, but these two individuals inaccurately entered their data; thus, they have been excluded from the estimating sample.

[11]Only the first session came close to being full. To help spread subjects between sessions, three or four subjects were rescheduled to the third session. Due to hardware difficulties in the first session, the second and third sessions started about 15 minutes late.

sessions. These communication effects, as well as self-selection into different sessions, could contribute to systematic differences between session participants.

TABLE 4. Estimation of Decision to Switch to Option B (Risk Aversion)

Coefficient	OLS Decision Num.	Tobit Decision Num.
Female	-0.540	-0.516
	(0.464)	(0.460)
Age	0.024	0.022
	(0.081)	(0.080)
Econometrics	0.927*	0.953**
	(0.471)	(0.467)
Session	0.485**	0.492**
	(0.224)	(0.222)
Constant	4.235**	4.268**
	(1.865)	(1.846)
Observations	65	65
R^2	0.137	.
R^2 Adj	0.0795	.
F	2.382	.
LL	-110.7	-112.5
df_r	60	61
N	65	65
Pseudo R^2		0.0393
Chi^2		9.208

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Excludes reversals and one individual who did not report age.

Table 4 provides coefficient estimates across individual characteristics predicting which decision an individual is likely to switch from the safe asset to the risky asset. A significant positive coefficient for a variable implies greater risk aversion. In the second column, I use a doubly truncated Tobit estimation since the dependent variable, can be no greater than 10 and no less than 1 (though no individual switched at decision 1).

I find that individuals who were enrolled in an econometrics class during the term of the experiment (Fall) tend to switch to option B approximately one decision later than comparable individuals who are not enrolled in econometrics. Note that this is not necessarily an effect of taking an econometrics class (though such an effect cannot be ruled out). This coefficient may suggest that individuals participating in econometrics classes during the fall term are more risk averse than those that take the class in a different (later) term or not at all. I find that individuals who participate in a later sessions appear to be more risk averse on average than individuals who participate in earlier sessions. Furthermore, I find that neither age nor gender play a significant role in risk aversion, although the age variance in this sample is very small.

Though some of the effects of individuals' characteristics on risk aversion are significant at the 5% level, these characteristics explain very little of the variation in the risk aversion parameter. The purpose of this risk aversion experiment was not specifically to establish the determinants of risk aversion; instead, the goal was to infer approximate individual coefficients of relative risk aversion to test whether risk aversion can help explain an individual's WTP to remove ambiguity. In the next section, I outline the experiment concerning WTP to remove ambiguity and test whether estimated relative risk aversion coefficients play a role in an individual's WTP. The findings from this section will be used to make inferences about how individuals appear to deal with ambiguity.

2.4 WTP to Reduce Ambiguity

Previous studies examining ambiguity and uncertainty using the composition of balls in an urn have focused on valuation differences between urns (Fox & Tversky 1995), or revealed preference choice, or paying to bet on drawing a ball from an urn with a known composition of balls versus an urn with an unknown composition (S.W. Becker & F.O. Brownson 1964). This last type of urn problem has become known as the Ellsberg urn– though Savage and Knight were instrumental in the use of this type of urn choice to explain vagueness of probabilities or ambiguity. A vast number of papers has been published using this framework, exploring topics such as subjective probability, subadditivity of subjective probabilities (D. Schmeidler 1989), comparative ignorance (Fox & Tversky 1995), and max-min approaches to rationalize decisions under uncertainty (Y. Halevy 2007). Since standard economic theory cannot explain many of the behaviors exhibited in these urn experiments, many different theories have attempted to rationalize actions observed in experimental settings. None of these studies have focused on how much individuals may be willing to pay to reduce the effects of ambiguity from the ambiguous urn.

Instead of the standard revealed preference experiment design, between an urn of unknown risks and known risks, I use a unique framework wherein an individual is forced into investing in an asset with ambiguous risks. Individuals are provided with the opportunity to pay to remove the ambiguity from an urn with an unknown composition of balls as opposed to being provided with the option to invest using an urn with a known composition of balls (or risks). This unique design allows one to measure the amount of ambiguity individuals are willing to pay to remove. Previous studies using the revealed preference method conclude that risk aversion is not necessarily a determinant of an individual's preference to bet on an urn with a known composition as opposed an unknown composition. Using the risk aversion measures discussed above, one is able to test the role that risk aversion may play in an individual's WTP to reduce the effects of ambiguity.

2.4.1 Experiment Design

Individuals were endowed with a budget w. They were then were given an asset, in the form of an urn filled with 10 balls. For this urn, they had to make an investment of their entire budget on pulling a Black or a Red ball out of the urn. If the ball that came out of the urn was the same color as the one they predicted, then they would double their investment; if the ball was a different color, then they would lose the entire amount invested. Subjects did not know the actual compostion of Black or Red balls in the urn, but they had to bet on either a Red or a Black ball being pulled from the urn.

However, each subject had the opportunity to sample the urn, at a cost of (c) per ball, to learn more about the composition of Black and Red balls in the urn. In this experiment, a subject had a budget, w = \$2, and to sample from the urn cost \$0.10 per ball (c =\$0.10). If a subject wanted to sample 3 balls it cost \$0.30. After choosing the number of balls that would be sampled, the colors of the (3) balls from the urn would be shown simultaneously to the subject before being placed back into the urn. A subject then had to bet the full amount left in his budget (\$1.70) on either Black or Red. After placing this bet on Black or Red, the computer drew one of the 10 balls from the urn at random. If the subject chose the color of the drawn ball correctly, his payoff was twice what he bet. In the example above, this amount is \$3.40. If the subject chose a different color ball than the one drawn, then he received a payoff of \$0.00 (i.e. he lost the full amount of his wager).

This experiment forced subjects to consider ambiguity in their decision making. Subjects were uncertain about the composition of balls in the urn and ambiguity arises from the subject's lack of information about the second order probability distribution dictating the mixtures of balls in the urns. Subjects were not informed of the random generating process used to create the composition of balls in the virtual urns, which was a symmetric binomial distribution. Thus, there was ambiguity as to which random generating process was used to fill the urns. If a subject chose to learn the entire composition of balls in the urn, then the bet still contained risk (unless all the balls in the urn were of one color). Even if all the balls were sampled, then subjects still had little idea as to what the distribution of possible mixtures might be. But the effect of "not knowing" the second order probability distribution (i.e. the effect of ambiguity) is removed by such sampling.

This experiment design allows us to observe the extent to which individuals are willing to pay to remove ambiguity. This experiment design is quite similar to individual investment decisions in financial markets. An individual may be willing to pay to get information about an asset (stock or mutual fund), which may inform him or her of some of the risks associated with investing in the asset, but there is still a great deal of unknowable uncertainty that remains about investing in such an asset.

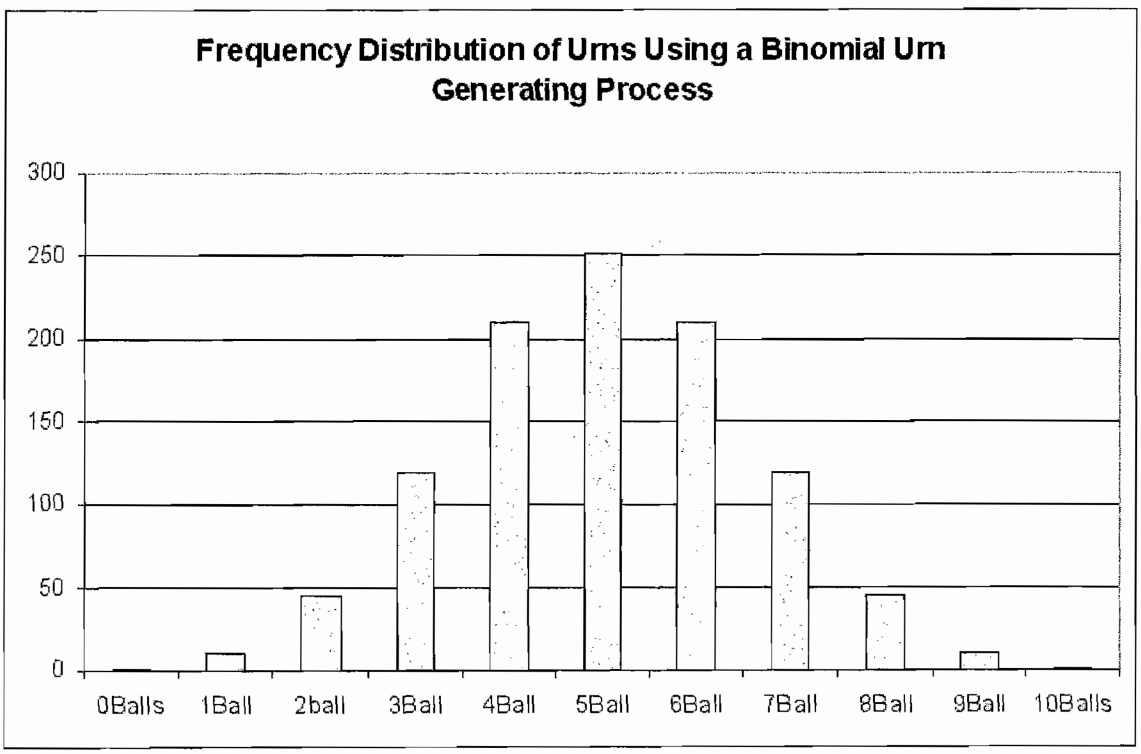
2.4.2 Maximization in the Urn Problem

The sampling of the urn, at first glance, should be a simple expected utility problem. In the classic Ellsberg example, if an individual has the choice between an urn filled with 5 Black and 5 Red balls, or an urn filled with 10 balls for which any mixture of Red and Black balls could be present, then that individual should be indifferent between betting on pulling out a certain color from either urn. Under expected utility theory, this indifference holds for any symmetric distribution; thus, so long as an individual does not believe that the urn was not generated in a manner that skews the composition towards a certain color then he or she should be indifferent. The innovation in this experiment is that there is no choice of urns; an individual must consider whether he can benefit from sampling rather than investing in an asset based upon symmetric odds of loss or gain. The benefit to sampling is based on an individual's prior beliefs about the possible distributions of different compositions of urns. This difference is explored below.

Consider an urn containing a mix of balls that is generated using a symmetric binomial process where the probability of a ball being Red is 0.5 (i.e. $p_{red} = 0.5$) and the probability of a ball being Black is the same (i.e. $p_{black} = 1 - p_{red} = 0.5$). If an urn has 10 balls then the number of unique possible urns is 1,024. The frequency distribution of such an composition generating process is shown in Figure 2 and the distribution approximates a normal distribution.[12] This type of urn-generating

[12]Historically, the binomial process was the first use of the normal distribution, introduced in Abraham de Moivre's book The Doctrine of Chances in 1733. Often in textbooks this approximation is shown as a consequence of the central limit theorem, since B(n, p) is a sum of n independent identically distributed Bernoulli variables with parameter p.

FIGURE 2. Frequency Distribution of Urns (Binomial)



process implies that the extreme outcomes of 10 Red balls or 10 Black balls are very unlikely. If, instead, the 11 possible compositions of urns are equally distributed, following a uniform distribution, then an urn filled with 10 Red balls is equally as likely as an urn filled with 5 Red balls. The difference between these two distributional assumptions can lead to very different optimal behaviors when it comes to making costly samples from the urn. The frequencies and probabilities are shown in Table 5 to highlight the probability differences between these two different distributions.

TABLE 5. Urn Compositions with Different Distributions

No. of Red Balls	Uniform Urns	Uniform Prob.	Binomial Urns	Binomial Prob.
0	1	0.091	1	0.001
1	1	0.091	10	0.01
2	1	0.091	45	0.044
3	1	0.091	120	0.117
4	1	0.091	210	0.205
5	1	0.091	252	0.246
6	1	0.091	210	0.205
7	1	0.091	120	0.117
8	1	0.091	45	0.044
9	1	0.091	10	0.01
10	1	0.091	1	0.001
Total	11	1	1024	1

There are many other possible distributions of the number of red balls, but for simplicity I limit this comparison to these two symmetric distributions. Initially, an individual has no reason to believe one distribution is more probable than the other. Thus, this experiment provides the opportunity to extract some notion of individuals' priors about the distribution of urns that they face.

In this experiment, the individual is assumed to maximize his or her expected utility by choosing to sample a certain number of balls from the urn prior to betting. This maximization problem is shown for individual i in Equation (2.4.1) where α is the dollar amount invested in the draw from the urn.

$$\max_{x\,\epsilon\{0,10\}} \sum_{j=1}^{11} U_i(z\alpha)\, P(z = 2|x, \Theta_i) \qquad s.t.\;\; \alpha + cx = w \tag{2.4.1}$$

The variable z provides the return if the ball. In this experiment $z = 2$ when the ball matches the one color that was bet on. The number of balls sampled in advance is represented by x and c is the cost of sampling each ball. Individual i's subjective ex-ante distribution for the return z depends on that individual's prior beliefs, Θ_i, about the distribution of balls and possible urns. The function $P(z = 2)$ represents the probability of the individual betting on a color and being correct. The expected utility depends on the individual's updated subjective distribution for return z after sampling a certain number of balls, x, from the urn. The risk aversion experiment task from above has shown that the group of individuals participating in this experiment is heterogeneous in terms of risk preferences; thus, this heterogeneity is represented by a different utility function form $U_i(\bullet)$ for each individual i.

The distributional assumptions concerning the probabilities for each of the eleven possible urns affect how many balls an individual should optimally sample. In this experiment the return on the urn can be one of two possibilities, $z \,\epsilon\, \{0, 2\}$ as the individual receives twice his wager if he is correct about the color of the ball pulled from the urn and receives zero if he is incorrect. I will assume that every individual's utility function is such that $U_i(0) = 0$. For simplicity, I will simplify this analysis of the maximization problem to focus only on expected values. The possibility exists that individuals' prior beliefs, in terms of urn distributions, are correlated with the type of utility function they have. In other words individual's risk preferences may be correlated with beliefs. In the following section, I explore the ex-ante expected values from sampling under two different distributions, the uniform distribution and the binomial (normal) distribution. In the subsequent section, I test whether individuals' estimated risk aversion parameters appear to influence how many balls are sampled from the urn. If the risk aversion parameters are positively correlated with sampling then this would suggest that ambiguity is processed by individuals in a similar manner as risk and uncertainty.

2.4.3 Expected Value Assuming Different Distributions

To examine the effects of different distributional priors regarding the space of possible urn compositions, consider a uniform distribution of urns compared to the binomial (normal) distribution. An individual is willing to sample the urn so long as the expected benefit from sampling is greater than the cost. In deciding how many balls to sample, an individual compares the net benefits of sampling differently or not sampling at all. An urn could be composed of 11 different possible combinations of Red and Black balls; sampling a single ball allows an individual to remove one possible urn from the 11 possible combinations. For example, if an individual observes a sample of a Red ball from the urn, then that individual has learned that this urn is not entirely composed of Black balls. In sampling a single ball an individual eliminates an urn that would be entirely composed of a single color. The benefit of sampling depends on how probable the extreme compositions of urns being filled with one color actually are. Consequently, fatter tails in the probability distribution of outcomes, such as the uniform distribution, makes sampling more beneficial. This occurs because each ball sampled removes a single possible composition; with a uniform distribution each possible combination of balls has an equal probability of occurring.

With a uniform distribution of urns, an individual has an ex-ante optimal number of balls they should sample. Given that not sampling the urn gives an individual a 50% chance of drawing a Red or Black ball. This probability exists so long as the individual's prior beliefs do not involve skewing the possible distributions towards one color or another, that is if the individual believes that one possible distribution is skewed towards Black balls then he also believes there is an equal likelihood of an equivalently skewed distribution towards Red balls. Furthermore, the possibility exists that if he or she draws an even number of balls then he or she has gained no benefit from sampling. This lack of benefit exists because an even ball sample can provide a symmetric number of Red balls to Black balls. Recall that in this experiment individuals are endowed with $2 and the cost for each ball sample is $0.10. After sampling the return to correctly predicting the color of the ball to be drawn is 100% and the return to being wrong about the color costs 100%. Table 6 provides a comparison of the ex-ante expected values of sampling between the two possible distributions that are being considered.

TABLE 6. Expected Values from Sampling

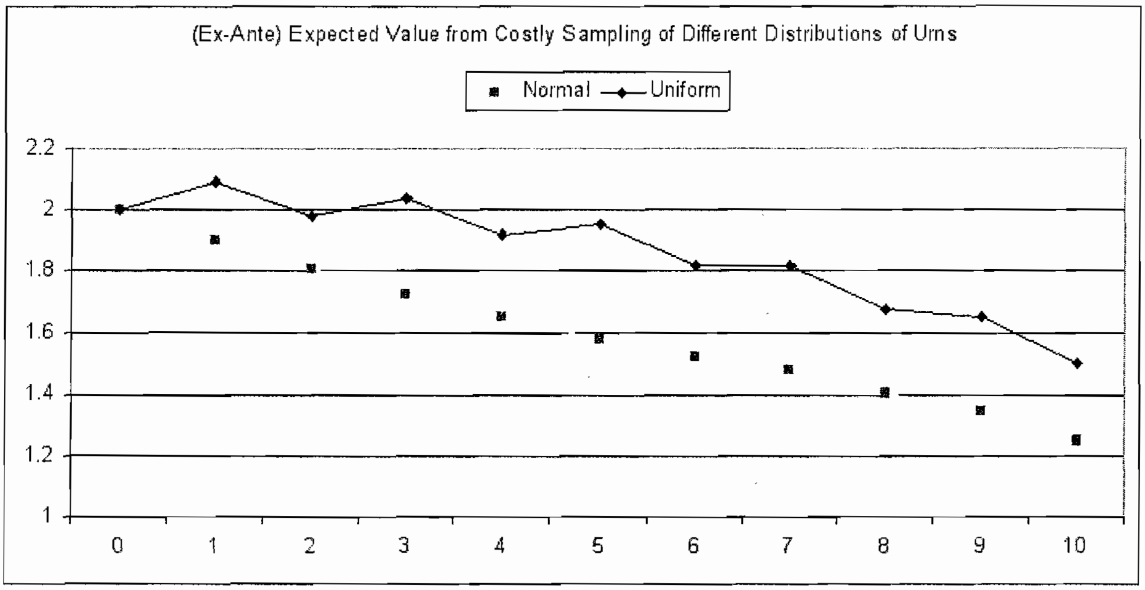
	Uniform		Binomial	
Ball Sample	Expected Value	Marginal △	Expected Value	Marginal △
0	2.00	0.00	2.00	0.00
1	2.09	0.09	1.90	-.10
2	1.98	-0.11	1.81	-0.09
3	2.04	0.06	1.73	-0.08
4	1.92	-0.12	1.65	-0.08
5	1.95	0.03	1.58	-0.07
6	1.82	-0.13	1.52	-0.06
7	1.82	0.00	1.48	-0.04
8	1.68	-0.14	1.41	-0.07
9	1.65	-0.03	1.35	-0.05
10	1.50	-0.15	1.25	-0.10

From Figure 3, one can observe the prior beliefs about the distribution of red and black balls in the urn play a pivotal role in deciding whether it is optimal to sample from the urn. An individual who has a prior belief that the set of possible compositions of balls in the urns was formed using a binomial method should not sample the urn at all. On the other hand, an individual who believes that the possible compositions of urns are uniformly distributed is justified in sampling one or three balls as opposed to none, since both these sampling strategies lead to higher expectedvalues. However an individual with such a prior belief would find it optimal to just sample one ball.[13]

The purpose of this analysis was to show that based on two different prior beliefs about the urns, an individual may be justified in sampling one, or three balls as compared to the option of not sampling any. Thus, this sheds new light on the

[13]Underlying this entire analysis is the assumption that an individual will bet according to the signal he or she receives. For example, if an individual samples three balls and observes Red, Red, Black then that individual will bet on Red.

FIGURE 3. Expected Value ($) from Sampling Urn by Sample Quantity



Ellsberg urn problem because if an individual was truly indifferent between an urn with an unknown composition and an urn with 5 Red balls and 5 Black balls then one would expect that sampling would bring no value to that individual. In the following section, the results from the urn experiment are analyzed. I look for systematic behavioral differences to investigate the possible cause of a large variance in urn sampling decisions.

2.4.4 Urn Experiment Results

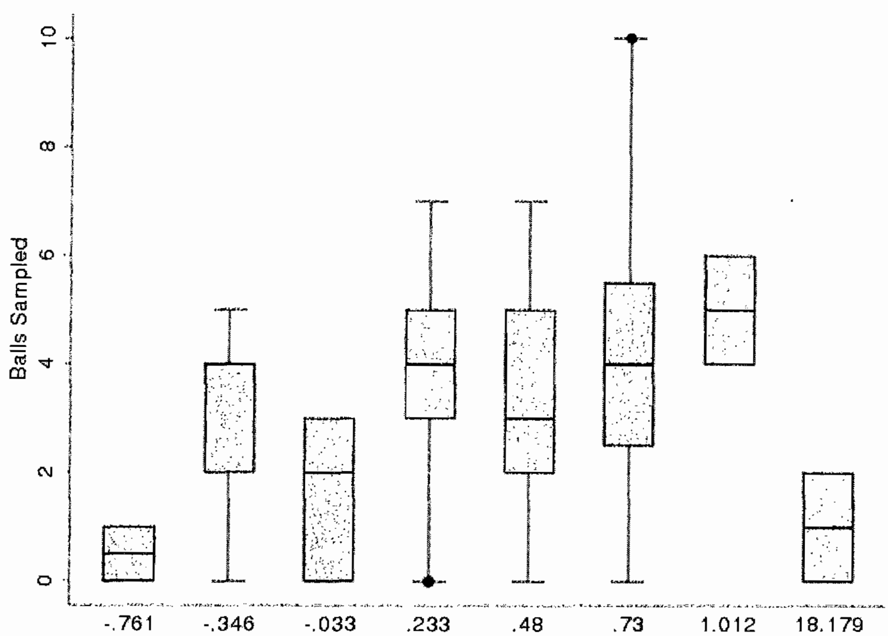
The urn experiment was run in three separate sessions and a total of 68 participants provided data that was usable for this analysis. There is a surprisingly large amount of variation in the number of balls sampled from the urn prior to betting on Red or Black. The frequency distribution in Table 7 shows that almost 20% of subjects chose not to sample the urn and over 66% of subjects chose to sample between between 2 and 5 balls.

TABLE 7. Frequency of Balls Sampled

Item	Number	Percent
0	13	19.1
1	1	1.5
2	7	10.3
3	12	17.6
4	15	22.1
5	11	16.2
6	5	7.4
7	2	2.9
8	1	1.5
10	1	1.5
Total	68	100.0

According to predictions about behavior concerning the Ellsberg Urn, the sampling of the urn should not be related to risk aversion because individuals will treat the urn as a compound lottery across all possible distributions. Based on the previous section that explored two different distributional assumptions, I hypothesize that individuals who are more risk averse will sample the urn more because their priors weight the tails of the distribution which in reality are the more extreme outcomes. Compared to a less risk averse individual, a more risk averse individual forms a prior distributional belief that the extreme outcomes of having an urn filled with all (or many) Red balls or all (or many) Black balls are more likely to occur than other compositions. To consider this, I use the CRRA values for individuals found from the risk aversion experiment that was discussed previously. Figure 4 shows the boxplots of balls sampled according to the individuals' CRRA. If one ignores the extreme form of risk aversion (CRRA=18.179), these boxplots suggest that there is a positive relationship between an individual's risk aversion coefficient and the number of balls he or she samples from the urn.

FIGURE 4. Number of Balls Sampled According to CRRA Estimates



To explore whether the estimated CRRA can help predict the number of balls an individual samples from the urn, I estimate both ordinary least squares (OLS) and Tobit regression models. These estimates are shown in Table 8. In these models, I include variables identifying individuals as being currently enrolled in an econometrics class and whether they are female. Due to the non-randomness of session attendance, I include a variable for session attended (1, 2, 3). I find that after including these control variables that the CRRA has a significant positive coefficient with respect to the number of balls sampled.

2.5 Discussion

These results suggest that risk aversion has an effect on how an individual deals with ambiguity. The Ellsberg urn problem is often given as an example of ambiguity where an individual is willing to bet on an urn with a known composition as opposed to one with an unknown composition. The results from this experiment suggest that risk aversion may play a role in how an individual approaches such an ambiguous situation. I caution that this does not imply that this risk aversion effect has a direct effect on ambiguity aversion. Instead, risk averse individuals may have very different prior beliefs about the distributions of urns compared to individuals who are less risk averse. I believe that there may be a positive correlation between risk aversion and the over-weighting of the tails of distributions. Thus, risk aversion may be linked to the subjective probabilities used for evaluating ambiguous outcomes by individuals.

In a previous section, it was shown that some sampling of the balls in the urn can be value maximizing. Sampling more than zero balls implies that individuals must believe they are increasing expected utility based on their distributional assumptions. The question becomes, do these individuals behave within a rational subjective utility framework? Do they choose a sample of balls that demonstrates plausible beliefs in terms of the beliefs about the probability of success from sampling? Based on the experiment performed, assuming CRRA utility function and symmetric beliefs

TABLE 8. Estimation for Urn Sampling

Coefficient	(1) OLS	(2) Tobit	(3) OLS	(4) Tobit	(5) OLS	(6) Tobit
CRRA	1.92**	2.32**	2.28***	2.75***	1.75**	2.09**
	(0.82)	(0.99)	(0.85)	(1.01)	(0.85)	(1.02)
Econometrics	-1.88**	-2.41***	-1.98***	-2.51***	-1.77**	-2.26**
	(0.72)	(0.88)	(0.71)	(0.87)	(0.74)	(0.90)
Session	-0.81**	-1.01**	-0.81**	-1.01**		
	(0.34)	(0.41)	(0.34)	(0.41)		
Female			1.10	1.29	1.09	1.27
			(0.73)	(0.87)	(0.76)	(0.91)
Constant	4.63***	4.72***	4.34***	4.39***	3.00***	2.73***
	(0.69)	(0.82)	(0.71)	(0.83)	(0.45)	(0.54)
Sigma						
Constant		2.38***		2.33***		2.45***
		(0.25)		(0.25)		(0.26)
Observations	62	62	62	62	62	62
R^2	0.184	.	0.215	.	0.136	.
R^2 Adj	0.141	.	0.160	.	0.0909	.
F	4.350	.	3.903	.	3.033	.
LL	-130.1	-127.5	-128.9	-126.4	-131.9	-129.5
df_r	58	59	57	58	58	59
N	62	62	62	62	62	62
Pseudo R^2		0.0485		0.0566		0.0339
Chi^2		13.00		15.17		9.091

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Excluded reversals in risk aversion experiment and subjects with extreme CRRA concerning outcomes, I am able to calculate a lower bound on individual's subjective beliefs. First, consider an urn with one ball that may be black or red.

2.5.0.1 1 Ball Example

The one ball case is an example of certainty equivalence using subjective prior beliefs, π. There is an urn with one ball, the ball is black or red, and the individual may sample the ball beforehand at a cost of c. An individual bets his net-budget, X, on pulling out a red or black ball. If the individual pulls out a ball of his predicted color, he doubles his bet. If he does not, he loses everything. The net-budget meets the constraint $X = W - c$. Where c is the cost of sampling the ball beforehand and W is the individual's initial wealth. In this one ball case the individual has a prior belief, π, on pulling out his preferred (if skewed beliefs) ball. If initially he has indifferent symmetric prior beliefs about outcomes, then $\pi = 0.5$.

The individual is faced with an ex-ante decision based on his priors. If this individual chooses to sample the urn, then the following inequality must hold for an expected utility maximizer:

$$U(2(W - c)) \geq \pi U(2W) \tag{2.5.1}$$

Assume a utility function as follows:

$$U(x) = \frac{x^{1-r}}{1 - r} \tag{2.5.2}$$

Then the inequality simplifies to:

$$(\frac{W-c}{W})^{1-r} \geq \pi \qquad (2.5.3)$$

Assuming that the above equality is strict and constant beliefs such that $\pi \in (0,1)$, then the willingness to pay to see a ball is greater for a more risk averse individual.

$$\frac{\partial c}{\partial r} = -\frac{1}{(1-r)^2} W \pi^{\frac{1}{1-r}} \ln \pi \geq 0 \qquad (2.5.4)$$

2.5.1 Derivation of Belief Bounds

Consider a sampling of balls that allows an individual to adjust his beliefs about successfully predicting which color of ball he will pull out of the urn. Beliefs, π, are a function of the number of balls an individual can sample, n. Assume that an individual gains a non-negative ex-ante benefit from sampling a ball.

$$\pi(n) \text{ such that } \pi(n+1) \geq \pi(n) \qquad (2.5.5)$$

In the experiment an individual chooses n such that $n \in \{0, 10\}$. The individual's utility maximization problem for choosing the optimum number of balls to sample is as follows:

$$\max_{n \in \{0,10\}} \pi(n) U(2(W - Cn)) \qquad (2.5.6)$$

Assume the same CRRA utility function as before. The optimal number of balls an individual should sample, $n*$, satisfies the following inequality:

$$\pi(n^*)U(2(W - Cn^*)) \geq \pi(0)U(2W) \tag{2.5.7}$$

$$\pi(n^*)(2(W - Cn^*))^{1-r} \geq \pi(0)(2W)^{1-r} \tag{2.5.8}$$

$$\Rightarrow \ln \pi(n^*) \geq \ln \pi(0) - (1-r)\ln(1 - \frac{Cn^*}{W}) \tag{2.5.9}$$

Based on the data gathered from the experiment and assuming that $\pi(0) = 0.5$ then the RHS of the last inequality can be used to extract a lower bound on beliefs of sampling. We can observe what is the lowest probability that an individual expected to gain from sampling the number of balls that he or she chose to sample. These belief bounds are based on the number of balls sampled and an individual's CRRA. For individuals with a CRRA that is greater than one, the belief bound decreases with a greater number of balls sampled. This decrease does not mean that individuals believe that the probability is not improved from sampling; instead this decrease implies the bound is decreasing. Thus, little information can be extrapolated regarding these individual's beliefs whose estimated CRRA is greater than 1.

Table 9 shows the number of individuals for each respective lower bound on beliefs found from sampling the urn. Individual choices made in regards to sampling the urn are shown in the row groupings and the estimated CRRAs of individuals are separated into columns. The most striking result this table provides is that no individuals behave in a manner that suggests irrational subjective beliefs. Over 65% of extracted lower bounds of beliefs are from 0.522 to 0.623. Thus, individuals believe

TABLE 9. Ex-ante Belief Bounds & Samples Taken by CRRA

	CRRA								
SamplesTaken	-.760	-.346	-.033	.233	.480	.730	1.012	18.179	Total
0	1	0	1	3	4	2	0	1	12
Belief	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	
2	0	0	1	0	3	1	0	1	6
Belief	0.602	0.576	0.557	0.542	0.528	0.514	0.499	0.08	
3	0	1	1	3	5	2	0	0	12
Belief	0.666	0.622	0.591	0.566	0.544	0.522	0.499	0.031	
4	0	2	0	7	1	3	1	0	14
Belief	0.741	0.675	0.630	0.593	0.562	0.531	0.499	0.011	
5	0	1	0	6	3	1	0	0	11
Belief	0.830	0.736	0.673	0.623	0.581	0.540	0.498	0.004	
6	0	0	0	1	2	1	1	0	5
Belief	0.937	0.808	0.723	0.657	0.602	0.551	0.498	0.001	
7	0	0	0	1	1	0	0	0	2
Belief	1.068	0.893	0.780	0.696	0.626	0.562	0.497	0.000	
8	0	0	0	0	0	1	0	0	1
Belief	1.229	0.994	0.848	0.740	0.652	0.574	0.497	0.000	
10	0	0	0	0	0	1	0	0	1
Belief	1.694	1.271	1.023	0.850	0.717	0.603	0.496	0.000	
Total	1	4	3	21	19	12	2	2	64

Individuals that had reversals in the risk aversion task have been removed.

that sampling from the urn can improve their probability of success by revealing something about the distribution of balls in the urn.

In many ways, this sampling of the urn is similar to purchasing information about an ambiguous or uncertain asset. Similar to investment markets, individuals seem to value information too much. In the case of the urn, the individuals in the experiment believe that sampling the urn can lead to a much greater improvement of subjective odds concerning success than it actually does. In fact, in extreme cases, such as a uniform distribution across urns, some sampling can be beneficial, but the greatest gain from sampling is at 1 ball. Not a single particpant chooses to sample only 1 ball. Thus, individuals must have very different distributions in mind prior to sampling. Or they need to have enough uncertainty removed from the choice problem (Red or Black) before making a decision.

In the urn experiment, subjects behaved such that sampling the urn was positively correlated with risk aversion. I have argued that sampling the urn is an attempt to reduce the effects of ambiguity associated with this urn problem. Prior to sampling, the urn with an unknown composition of balls, can be considered an ambiguous asset. There are a number of different possible distributions that an individual may have in mind when speculating about the composition of balls in the urn, but these distributions are not known to the individual ex-ante. Thus, to remove the effects of all the possible distributions, and to gain some certainty as to what the composition of balls in the urn may be, an individual chooses how much of this ambiguity to remove. Through this experimental design we are able to observe (under expected utility maximizing assumptions) the bounds on beliefs regarding improvements on outcomes from sampling that individuals must have to make certain sampling choices. This could be useful in helping predict optimism or pessimism towards other perceived outcomes that may affect individuals.

2.6 Concluding Remarks

The purpose of this study has been to test whether individuals are willing to pay to remove the effects of ambiguity. It can be shown, in theory, that individuals can find it value maximizing to sample the urn and remove some ambiguity depending on their prior beliefs about the distribution of possible risks and outcomes and the cost of sampling. I find that many individuals sample the urn to an extent which suggests that their prior beliefs likely do not follow a uniform or normal distribution. The results do not imply that individuals behave irrationally, however, based on the lower bounds derived for individuals' prior beliefs, they still behave rationally, though their distributional priors maybe difficult to infer. This series of experiments showed that using this experimental design can provide at least some notion of what those prior beliefs may look like for individuals.

Risk aversion should not influence decisions to remove the effects of ambiguity at a cost, but individuals who exhibit greater risk aversion tend to sample more balls from the urn. This result suggests that risk averse individuals may demand more information even if that information is not linked with risk, but is instead linked with ambiguity. The urn sampling experiment provides a measure different from risk aversion. This other measure–this WTP to remove ambiguity–may be useful in predicting behaviors that a simple measure of risk aversion cannot, or does so quite differently. The WTP to remove ambiguity and risk aversion coefficients may measure very different things because the combined measures provide different bounds on individual beliefs.

This study demonstrates support for a fairly common belief, that risk and ambiguity are different from each other, though they may be correlated. The experiment design used here, that isolates how individuals wrestle with ambiguity, has provided a measure that may have predictive qualities even in a strategic setting. But these results, though interesting and powerful, do raise more questions. Individual aversion towards ambiguity is likely to matter for other behaviors, but what is the extent of these effects? Is there a way to translate such a measure of WTP to remove ambiguity to other settings? In the end this study suggests that ambiguity aversion affects behaviors in a different manner than risk and it provides a manner by which one can proxy for individual ambiguity aversion in laboratory settings by using the WTP to remove ambiguity task design.

CHAPTER III

DIFFERENCES IN COMPETITIVE CHOICES: PERFORMANCE FEEDBACK, GENDER AND HORMONAL PHASE.

3.1 Introduction

Labor markets include a variety of different forms of compensation or competitive environments. Firms may pay workers a flat rate, or instead they may incentivize workers using a piece rate or through an economic contest, such as a rank-order tournament (E.P. Lazear & S. Rosen 1981). These different forms of compensation and contests are used in a variety of different markets with the motive being to increase effort, output or performance or to sort high and low ability workers whose types are unobservable.

Ability is not the only factor to influence this sorting. Females are significantly less likely to choose a competitive tournament when given the choice between a tournament or a piece rate form of compensation, even with no performance differences (Muriel Niederle & Lise Vesterlund 2007). It has been suggested that both risk aversion and preference differences for competition are partly responsible for these differences between the sexes, but the role that relative performance feedback may have on these gender differences has not been explored. In this paper, I attempt to answer whether differences in competitive environment selection are affected by fully informing agents about their relative performance compared to other agents they may compete against.

To examine effects of relative performance feedback, I use an economic experiment with two specific treatments to analyze the selection of competitive environments. In the uninformed treatment, with no relative performance feedback, subjects receive information only about their absolute performance from a previous treatment in a real effort task. They do not receive any information about how well any other participants performed. This uninformed treatment is followed by an informed treatment, where subjects receive information about the previous performance of all potential competitors along with their own performance. In both treatments, subjects make choices across the same set of competitive environments; thus, any differences in self selection between treatments may be attributed to the role that relative performance feedback has in these decisions. I find that relative performance feedback removes any systematic gender differences in the selection of competitive environments.

The possible biological mechanisms behind these behavioral differences between the sexes have only begun to be explored in economics.[1] Previous studies have concluded that females and males have differences in preferences and behaviors for competitive environments (Niederle & Vesterlund 2007), risk (P.J. Grossman & C. Eckel 2003), and investment behavior (B.M. Barber & T. Odean 2001, G. Charness & Uri Gneezy 2007). A possible biological basis for these differences comes in the form of the hormonal differences between females and males. Females' hormones,

[1] For a review combining findings in economics and psychology see Matsushita, Baldo, Martin and Da Silva (2007).

specifically steroid hormones, fluctuate a great deal and in a predictable manner across the menstrual cycle.

The economic impacts of hormonal effects caused by the menstrual cycle are potentially significant, but have only been studied in a few domains. For example, Ichino and Moretti (2009) found that female worker absenteeism may be partly a function of menstrual cycles. In this study, by scheduling females to participate in two sessions, during both a low and high-hormone phase, I test whether hormonal differences in females are related to competitive environment selection. I find that females participating during the low-hormone phase are less likely to enter competitive environments than females in a non-low-hormone phase of the menstrual cycle in the treatment with no relative performance feedback. But as with the gender differences, these selection differences are removed with relative performance feedback.

In the following section, I review the previous literature concerning competitive environments and gender differences. I also review the literature and provide rationale for why hormonal fluctuations may play a relevant role in competitive environment selection for females. In section 3.3, I explain the experiment design and the subsequent sections follow with results, a discussion, and ending with concluding remarks.

3.2 Previous Literature

In examining the previous literature, I separate the two main areas of study: competition and gender differences, and the effects of hormones on behavior. The first section focuses on previous literature concerning competitive environments and gender differences. This section is then followed by the literature on the effects of hormones in economic decision making.

3.2.1 Competition and Gender Differences

The corporate ladder can be considered a tournament where a number of individuals compete for promotion based on the results of individual performance. 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. This underrepresentation of females in executive positions may be partially explained by the roles that females have in the traditional family with the raising of children affecting their career choices and human capital investments (S.W. Polachek 1981). Part of this underrepresentation may be caused by a preference by females to receive piece rate compensation. Jirjahn and Stephan (2004) find that the attractiveness of piece rate schemes for females is likely caused by less wage discrimination in such a setting 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 (C. Brown 1990).

The disproportionate number of females in high ranking executive positions may also be a result of preference differences for competitive environments, or lower performance gains for females from participating in tournament settings when compared to males. A few other possible explanations exist for gender differences in competitive environment selection. Tournament settings are more risky than piece rate settings, so the difference in self selection may stem from gender differences in risk aversion. Another possibility is that females and males evaluate expected values differently when the distributions of outcomes or potential competitors are ambiguous.

Niederle and Vesterlund (2007) found, when given a choice between a piece rate or a winner-take-all tournament compensation scheme, that females overwhelmingly choose the piece rate while males choose the tournament. They infer that this gender difference in selection is driven by men being overconfident and by differences in preferences for competition between females and males. They find that the gender differences for competition still remain even when they control for confidence and risk aversion. Datta Gupta, Poulsen and Villeval (2005) get similar results in an experiment examining the effects of gender composition in tournaments. In another experiment, Dohmen and Falk (2007) find that females are less likely to choose variable pay schemes such as tournaments and piece rates when given the alternative choice of a fixed rate for their time. All these experiments used a similar protocol where subjects were given their absolute performance, but were never informed of their relative standing within the group. These economic experiments have been interpreted as meaning that gender differences in self selection are derived from a preference difference where females seem to have a greater distaste for competition than males. One should note that socialization may contribute to gender differences in the selection of competitive environments as Gneezy et al. (2009) find that in a matrilineal society women are more likely to compete than men.

One reason for these choice differences could be from a performance difference between females and males. Gneezy, Niederle, and Rustichini (2003) find that females see lower performance gains from participating in competitive environments. In observing children's performance in running races, Gneezy and Rustichini (2004) also find that competition increases performance of boys, but not girls. In an experimental setting, Gneezy, Niederle, and Rustichini (2003) find that for a mixed-gender competitive environment, males have significant performance increases when an environment is made more competitive and 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 competitive environment decision. These results suggest that performance in competitive environments is different for males and females and that the gender composition of groups may play a role in performance gains from competition, as well as in the selection of competitive environments. Both performance differences and gender composition effects may help explain the underrepresentation of females in the corporate business world.

In another study, Niederle, Seagal and Vesterlund (2009) replicate previous findings of gender differences for competitive environment selection, and then examine the role of confidence and an affirmative action type of policy on these selection differences. They find that while affirmative action policies change the composition of the applicant pool, the overall number of high-performing participants is not substantially affected. Thus, due to high-performing women coming in at the cost of high-performing men, the performance costs of selecting women over men by affirmative action policies may be offset by these selection behaviors by high ability individuals. These results suggest that affirmative action policies may remove gender differences in the selection of competitive environments, and may not be costly, so long as there are no ability differences between males and females.

Systematic differences between genders in risk aversion may also contribute to differences in participation in competitive environments. Some studies find that females are more risk averse than males though results are inconsistent in laboratory settings (Grossman & Eckel 2003). These gender differences are not entirely robust because subjects from non-western cultures, and children, appear to not exhibit differences in risk preferences between the sexes (Charness and Gneezy 2007, Harbaugh et al. 2002). Risk aversion is a significant factor when making a decision to enter a competitive environment; however, the competition studies mentioned above control for risk aversion effects and gender differences still remain.

Another factor that may contribute to gender differences in tournament selection may be ambiguity aversion. In the experiments that found gender differences in competitive environment selection, individuals had little information about the performance distribution that they must consider in calculating the probability of winning the tournament. But ambiguity aversion differences have not been found to occur systematically between the genders. Moore and Eckel (2003) find that females are more ambiguity averse for specific contexts and domains. On the other hand, Borghans et al. (2009) find that males are initially more ambiguity averse than females, but as ambiguity increases, males and females behave similarly.

Previous studies have found gender differences in the selection of competitive environments. The possibility exists that these gender differences in competitive environment selection are driven by biological factors. In the following section, I explore why hormones may also play a pivotal role in these decisions.

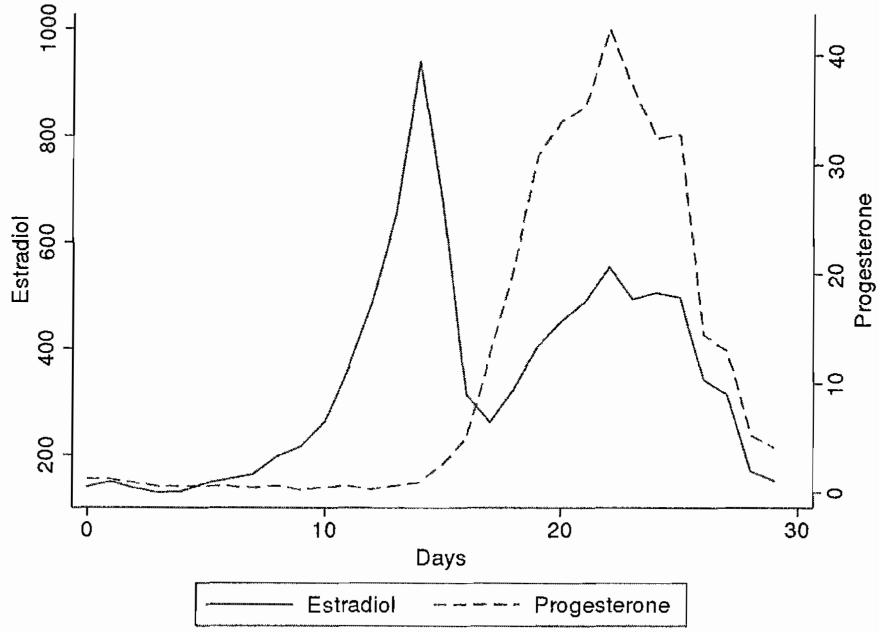
3.2.2 Hormones: Why might they matter?

A biological basis for differences in preferences and behaviors between the sexes may be due to hormonal differences between men and women. In addition to differences in hormone levels, females experience large and fairly predictable steroid hormonal fluctuations across the menstrual cycle. Steroid hormones have been found to matter for behaviors and economic outcomes in other contexts. Testosterone levels of financial traders in the morning can predict profitability through the rest of the day. Also, cortisol levels in these same traders were found to rise with increased volatility in their market returns (JM Coates & J. Herbert 2008). Testosterone levels are correlated with behaviors in economic experiments such as offers and acceptances in ultimatum games (T.C. Burnham 2007). Oxytocin has been shown to encourage generosity levels and trust in individuals (Kosfeld et al. 2005, Zak et al. 2005). Furthermore, through intranasal administration of oxytocin combined with fMRI scans, it was found that oxytocin reduces activation in specific areas of the brain related to fear processing and information feedback response (Baumgartner et al. 2008). Estradiol can increase power motivation–a preference for having an impact or dominance over individuals–in females suggesting that it may affect competitive appetites for females (S.J. Stanton & O.C. Schultheiss 2007).

Premenopausal females not using hormonal contraceptives experience significant fluctuations in hormones. Estrogen, progesterone, follicle-stimulating hormone (FSH), and luteinizing hormone (LH) all have a consistent pattern for normal cycling females. Estrogen and progesterone have received most of the attention in studies examining neuroendocrinological, psychological and behavioral effects. As shown in Figure 5, both estrogen and progesterone 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. Estrogen rises quickly and spikes just prior to ovulation. This rise is referred to as the follicular spike as it occurs during the follicular phase, which is just prior to ovulation. After ovulation (approximately day 14 in the graph), progesterone levels spike in the latter half of the female menstrual cycle, which is called the luteal phase. During the luteal phase, females who ovulate experience heightened levels in both estrogen and progesterone. This second spike in both hormones may be referred to as the luteal spike or high-hormone phase (Stricker et al. 2006, Speroff and Fritz 2005).

Hormonal contraceptives cause major changes in the hormonal fluctuations that occur across the natural menstrual cycle. Females using a hormonal contraceptive experience suppression of endogenous hormone production when in the active phase of their contraceptive regimen (L. Speroff & M.A. Fritz 2005). Both progesterone and estrogen levels remain fairly constant as the body receives a dose of these hormones exogenously (Aden et al. 1998). During the placebo or non-active phase of the contraceptive regimen there are no exogenous hormones being provided to the body; this exogenous lowering of both progesterone and estrogen leads to *withdrawal bleeding*

FIGURE 5. Hormones in Normal Cycling Females



Estradiol measured in pmol/L and Progesterone in nmol/L.
These are median values from Stricker et al. (2006)

caused by the withdrawal of exogenous hormones (Speroff & Fritz 2005). Interestingly, there is no biological or medical necessity to induce this withdrawal bleeding (FD Anderson & H. Hait 2003). This is made apparent in the following quote from one of the leading textbooks on clinical gynecologic endocrinology.

> *Monthly bleeding, periodic bleeding, or no bleeding–this is an individual woman's choice (Speroff & Fritz 2005, 908).*

It is believed that keeping withdrawal bleeding as part of the hormonal contraceptive regimen was a marketing ploy to make the birth control pill seem more socially acceptable (E.M. Coutinho & S.J. Segal 1999). Thus, females could entirely avoid hormonal fluctuations by sustained contraceptive use. Indeed, some forms of contraception do ensure that this occurs (Anderson & Hait 2003). If the decrease in hormones affects females in an costly manner then one could expect that there is some demand to prolong contraceptive use among the female population.

Neuroendocrinology has demonstrated the existence of hormonal effects on brain activity. Results show that major depression may be linked to reduced density of hydroxytryptamine (5-HT), also know as serotonin, binding sites (Malison et al. 1998). By injecting estradiol in rats, Fink et al. (1996) find that estrogen stimulates an increase in the density of 5-HT binding sites in certain areas of the brain. Injections of estradiol significantly increase the density of binding sites in the anterior cingulate cortex (24%), anterior frontal cortex (41%) and the nucleus accumbens (12%). These areas of the brain have been variously linked with mood, memory, and the anticipation and receipt of monetary rewards (Fink et al. 1996, McEwen 2002, Bethea et al. 2002, Platt and Huettel 2008). Progesterone has been shown to inhibit neurotransmission, and as a result it may decrease anxiety and increase sedation (E.L. Vliet 2001). Other research suggests that progesterone may decrease the degradation of 5-HT (Bethea et al. 2002).

Hormonal fluctuations occur across the female menstrual cycle and contribute to premenstrual syndrome effects, which may have significant economic consequences. Ichino and Moretti (2009) use detailed employee attendance data from a large Italian bank and find that absences for females below the age of 45 tend to occur according to a 28 day cycle. These 28-day cycle absences explain about one-third of the gender gap in employment absences at the firm. The female menstrual cycle is approximately 28 days and they focus on females below the age of 45, who are more likely to be premenopausal. Ichino and Moretti's result provides support for hormonal fluctuations of the menstrual cycle having significant economic effects.

Using economic experiments, Chen et al. (2005) explore hormonal differences as determinants of economic behaviors. They focus on behavioral differences between males and females in first-price auctions. They find that males and females behave differently only when females are in a phase of the menstrual cycle that provides heightened levels of estrogen. In contrast, Pearson and Schipper (2009) find that women bid higher than men only during the menstrual and premenstrual phases of the cycle when estrogen levels are lower.

Not all economic studies have found support for hormonal effects being significant in economic decision making. Zethraeus et al. (2009) examine 200 post-menopausal women in a double-blind study. Participants were given either estradiol (2 mg), testosterone (40 mg) or a placebo daily for a four week period. Then they participated in an economic experiment session that included a variety of different tasks looking at risk aversion, altruism, reciprocal fairness, trust and trustworthiness. No significant differences were found when comparing the behaviors between the three different treatment groups of females.

In the following section, I explain the experimental design used to examine both the effects of relative information and hormones on competitive environment selection.

3.3 Experiment Design

Previous experimental studies have concluded that preference differences for competition may be driving differences in selection for competitive environments by males and females. I use an experiment design to test whether preference differences are the main source of these gender differences. Use of a partial information treatment and a full information treatment creates an opportunity to observe the role that relative information may play in the selection of competitive environments.

Furthermore, no economic research has been done to examine the possible effects of hormonal fluctuations across the menstrual cycle and competitive environment selection. Therefore, the experiment design includes specific recruiting and scheduling to attempt to have subjects participate in two sessions; for females these sessions were to occur during both a low and high hormonal phase of the menstrual cycle. I also include females taking hormonal contraceptives since females may alter hormonal fluctuations by taking a hormonal contraceptive. For females using hormonal contraception, during the placebo phase of the hormonal contraception regimen, their hormone levels are the same as normal cycling females in the low-hormone phase of their cycle. This design allows for the comparison of behaviors between females in the low phase to those not in the low phase, while also comparing the behavior of male participants.

In this study, most subjects participated in two separate competitive task sessions. One session involved a word task and the other session involved a math task. There were multiple reasons to try to have subjects participate in two sessions.

The first reason was to examine whether females select competitive environments differently conditional on being in a high or low phase during their menstrual cycle. The second reason behind the two sessions and two separate types of tasks was to observe whether gender differences for competitive environments exist regardless of the type of task. Two sessions of participation also provided the opportunity to observe the stability of gender differences for similar competitive environments.

In experiment sessions, participants performed math and word based tasks requiring the exertion of real effort in five different treatments. In the first treatment, participants performed the task under a piece rate compensation scheme, which was non-competitive, since it was entirely dependent on the individual's performance in the task. In the second treatment, participants performed the task under a tournament scheme, they were randomly assigned to a winner-take-all tournament with a size of 2, 4 or 6 total competitors. This second treatment provided participants with experience in a randomly chosen competitive environment in which their pay depended on their own performance as well as the performance of others. In the third treatment, participants performed the task under a group-pay treatment. This treatment randomly paired participants with another participant and their total group production was split evenly between the two group members. This third treatment could be considered the least competitive since free-riding was a possibility for participants. They could be paid more than the non-competitive piece rate scheme so long as their partner's output was greater than their own.

These first three treatments were designed to provide participants with some experience in all of the possible compensation schemes. In the final two treatments, they were able to choose between the piece rate, group pay, or any of the tournament compensation schemes. In each of the first three treatments, participants were given only their absolute performance, they were not informed of how productive other participants in the session had been, or if they won or lost a tournament. In the fourth treatment participants could choose between all of the possible compensation schemes. These choices consisted of group pay, individual piece rate, or a winner-take-all tournament consisting of 2, 4, or 6 total competitors. In the fourth treatment, subjects were not informed of their relative performance compared to others in the session.

In the fifth treatment, participants were first informed of how all individuals in the session performed in the first treatment. They then chose between all the possible compensation schemes for their performance in the fifth treatment. The fifth treatment is used to examine whether gender differences that were previously found in similar competition experiments (Niederle & Vesterlund 2007, N.D. Gupta, A. Poulsen & M. Villeval 2005) were a result of preference differences, risk aversion, overconfidence or different approaches to information processing between genders. Providing complete relative performance information allowed participants to update their beliefs accurately. If gender differences remain in this treatment then such results provide support for the conclusions from previous experiments that observed gender differences in selection for competitive environments are a result of preference differences for competitive environments and risk.

3.3.1 Tasks

Two different types of tasks were used for this study, a math-based task and a word-based task. The math-based task was similar to the one used by Niederle and Vesterlund (2007). In this task, participants are asked to add four randomized two-digit numbers and complete as many of these summations as possible in a period of four minutes. Equations are presented to participants on a computer screen and they simply type in the answer and press the *Enter* key or click a *Submit* button on the screen. After each submission participants are promptly shown the next equation to solve. On the screen, the equations look similar to the following:

$$12 + 57 + 48 + 52 = \tag{3.3.1}$$

In the experiment, a sheet of paper and a pencil were provided to all participants to use for this task, but no other form of assistance was provided.

The word-based task was similar to one used by Günther et al. (2008). 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 as duplicate answers are counted as being 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 so long as these words still begin with the appropriate letter. In the experiment, participants were informed of the rules in advance of beginning the task. All participants were informed that everybody in the same session and same treatment received the same letter; thus, providing them with a task of equivalent difficulty for each treatment.

A select group of letters was used for this study to limit the variation of difficulty between treatments and sessions. The word list used for grading words is a common English word list used by open source word processors.[2] The letters chosen for the study had between 2.7% to 3.8% of all words in the word list beginning with these letters. The letters used are listed below with the percentage of words beginning with the letter in parentheses:

E	(3.8%)	F	(3.3%)	G	(3.0%)	H	(3.7%)
I	(3.6%)	L	(2.7%)	N	(3.5%)	O	(3.5%)

3.3.2 Compensation schemes

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. To mitigate differences in the payoffs between math and verbal based sessions, the payouts were slightly different for word tasks compared to math tasks. Payoffs for both the math and verbal tasks were calculated in a similar manner though the base rate, in the form of the piece rate (w) was different for word formation tasks ($w_w = \$0.25$) and math addition tasks ($w_m = \$0.50$).

[2]Spell Checking Oriented Word Lists (SCOWL), Revision 6, August 10, 2004 by Kevin Atkinson.

Suppose W_i is the total payoff from a treatment for individual i and y_i is the output of the individual and y_{-i} is the output of the other individual of interest that may have affect the payoff for individual i. In a binary group pay situation, individual $-i$ is the group member that is paired with individual i. For group pay, the payoff for individual i and symmetrically for individual $-i$ is:

$$W_i = \frac{w_k\left(y_i + y_{-i}\right)}{2} \tag{3.3.2}$$

In a tournament situation, if an individual has the best performance in his tournament 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, in terms of best performance, the individual receives a fraction of the tournament winnings based on the number of individuals he tied with. One should note that individuals were not informed about whether they won or lost a tournament until all five treatments were complete.

Individuals were informed that they could be randomly grouped with individuals that did not necessarily choose the same compensation option. This design creates a greater incentive for high ability individuals to choose a more competitive environment as there is a positive probability that they may compete against lower ability individuals. As well, this design creates an incentive for low ability individuals to choose group pay as there is a positive probability that they may be matched with a high ability individual; thus, increasing their expected payoffs.

3.4 Results

Experiment sessions took place in a computer lab at a large public university. Potential participants completed a screening survey online and were then scheduled for the first of two sessions. One of the two sessions involved a math-based task and the other session involved a word-based task.

TABLE 10. Summary of Attendees

Variable	Mean	Std. Dev.	Min.	Max.	N
Age	20.5	2.81	18	33	344
Years PS	2.19	1.48	0	6	343
GPA	3.28	0.47	2	4.22	343
Live Independently	0.82	0.38	0	1	345
Word task	0.48	0.5	0	1	345
Session Size	14.54	4.15	7	21	345
Second session	0.37	0.48	0	1	345
Female	0.5	0.5	0	1	345
Low Phase	0.14	0.34	0	1	345
Psych meds	0.08	0.28	0	1	343

The majority of participants were university students; the summary statistics for this standard student sample are in Table 10. The average size of the 26 experiment sessions was 14.5 participants with a standard deviation of 4.15. The word task was used in 12 of the sessions while 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.[3] Approximately half the sample consists of females (172) and 37% (126) of the sample consists of individuals that attended a second session. Based on self

[3] One individual (female) was removed from the data due to a lack of understanding of the stimuli.

reported data, 47 of the individual participants were females in the low-hormone phase of their menstrual cycle.

TABLE 11. Individual and Session Characteristics

Sex	Age	Years PS	GPA	Indep.	Psych	Word	Size	Sess. 2
Male	20.69	2.14	3.24	0.84	0.10	0.47	13.79	0.37
Female	20.32	2.24	3.32	0.81	0.07	0.48	15.30	0.36
Total	20.50	2.19	3.28	0.82	0.08	0.48	14.54	0.37

Table 11 shows that the males and females who participated were not noticeably different from each other. The one exception was that the average female participated in a session that was approximately larger by 1.5 individuals. In sessions, the female to male ratio ranged from 0.3 to 2.3. On average the female to male ratio per session was 1.01. Thus, all sessions had some degree of gender mix and on average this gender mix was approximately 1 to 1.

Sessions took place three to four times a week and were held in the morning. Sessions took slightly less than an hour, including approximately 10 minutes at the beginning of the session that were used for participants to wait together in a foyer. The reason for this initial wait period was to allow 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 membership were also anonymous because subjects were unable to know with whom they were paired or with whom they were competing against.

TABLE 12. Payouts by Task Type

Type	mean	sd	min	max	N
Math	7.38	11.31	0	84	180
Word	15.01	18.90	0	111	165
Total	11.03	15.86	0	111	345

The summary statistics for payouts for task performance are shown in Table 12. These payouts were based on one randomly chosen treatment and do not include the participation fee given to all participants.[4] The word task paid substantially more than the math task. This difference was mainly due to the word task being substantially easier than the math task, despite the different wage used.

Participants who attended two sessions were subsequently 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 try to avoid endogeneity with the competition task performance. 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. Thus, an individual who participated in two sessions and the risk aversion task received an average payout of $38.95 for approximately 2 hours of time.

[4]The participation fee was $5 or $10 depending on the individual.

3.4.1 Task Performance

Each individual (in both word and math task sessions) participated in five different treatments of tasks. All participants in the session started the tasks at the same time as everyone else and had four minutes to complete the task. 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 tourney).

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

TABLE 13. Performance Across 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

Table 13 provides mean performance for each treatment for both math and word of tasks. According to the mean values, some learning seems to occur in both the math and the word tasks. This learning appears to be limited to the first three treatments. Table 13 also provides mean performance levels for both males and females for math and word tasks across the five treatments. The gender difference for performance in the math task is only significant in Treatment 1 at the 10% level.[5] In the word task the differences in performance are not statistically significant for any of the treatments. Thus, there are no noticeable performance differences between males and females after the first treatment.[6]

3.4.2 Performance

The main focus of this study is on the compensation choices that individuals made in the experiment; however, these choices may have been affected by performance differences. This section focuses on task performance for the different treatments. To consider how individuals are affected by the different incentives of each type of compensation, I focus on the performance of individuals in the first three treatments. In these first three treatments, individuals had no choice over the type of compensation they received for their efforts; thus, the performance effects from different compensation environments are exogenously determined.

According to the theory of piece rates and tournaments, one would expect greater effort for a higher piece rate. Similarly, an individual of higher ability and higher probability of winning should increase effort in a tournament. As the tournament gets larger and more competitive, one would expect that individuals would increase effort or set their effort levels to zero. Before considering the effects of

[5] 1-tailed t-test

[6] The analysis of the performance effects that occur from learning and different competitive settings are explored in Section 3.4.2.

tournament size on effort, I first focus on possible order effects and gender differences between treatments.

The regressions in Table 14 are used to consider gender differences in performance, learning effects and the incentive effects of increasing tournament size. The performance in the word task, but not the math task, is highly correlated with the GPA of participants. Regression estimates for both word and math show an order effect suggesting that subjects are learning in the first three treatments. Regression 1 in Table 14 shows that the tournament size has a statistically significant positive effect on individual performance in the math task. Increasing the competitiveness of a compensation environment from the piece rate to a tournament size of 6 should increase performance of an individual by 0.65 problems in a four-minute task. This is an increase of 5.7% for the average individual. In columns 2, 4, and 5, categorical variables are used to investigate whether tournament size is actually leading to the increase or whether just competing against someone in a tournament of any size leads to performance increases.

In the second column in Table 14, categorical variables were used for each of the possible competitive environments, for group pay and for tournaments. To test whether the tournament size matters a separate dummy variable, Tourney (ts>1) was used to identify if an individual had to compete against someone else. This categorization created a separate baseline for tournaments consisting of six individuals. Once controlling for tournament competition it was found that the size of the tournament does not matter and that group pay performance is not significantly different from the piece rate environment in the math task. There is also a positive

TABLE 14. Performance with No Choice

VARIABLES	(1) Math	(2) Math	(3) Word	(4) Word	(5) Low Math
Task Order	0.88	0.82	3.49	3.50	1.10
	(0.09)	(0.09)	(0.39)	(0.38)	(0.11)
	(***)	(***)	(***)	(***)	(***)
Tourney Size	0.13		0.04		
	(0.04)		(0.15)		
	(***)		()		
Tourney(ts>1)		0.63		-0.16	1.18
		(0.27)		(1.02)	(0.31)
		(**)		()	(***)
Tourney Size=2		-0.07		2.08	-0.24
		(0.38)		(1.49)	(0.42)
		()		()	()
Tourney Size=4		-0.21		0.22	-0.73
		(0.37)		(1.40)	(0.45)
		()		()	()
Female	-0.45	-0.46	2.31	2.29	0.44
	(0.59)	(0.59)	(1.87)	(1.88)	(0.39)
	()	()	()	()	()
Years PS	-0.50	-0.50	0.16	0.15	0.17
	(0.29)	(0.30)	(0.87)	(0.87)	(0.23)
	(*)	(*)	()	()	()
Age	0.40	0.40	0.33	0.34	-0.07
	(0.15)	(0.15)	(0.49)	(0.49)	(0.14)
	(***)	(***)	()	()	()
GPA	-0.26	-0.26	6.60	6.57	0.32
	(0.63)	(0.64)	(1.94)	(1.95)	(0.41)
	()	()	(***)	(***)	()
Constant	3.22	3.39	4.11	0.00	6.43
	(3.42)	(3.43)	(11.57)	(0.00)	(2.58)
	()	()	()	()	(**)
Letter Controls	No	No	Yes	Yes	No
Observations	534	534	492	492	303
Number of id	178	178	164	164	101
R-sq	0.0745	0.0721	0.367	0.368	0.196
chi2	102.5	104.5	565.4	570.9	136.0

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

effect for age as the average effect of a year of life leads to an increase in performance in the math task of 0.5 problems, though this may be offset by further post-secondary schooling. It is worth noting that overall the competitive environment and individual characteristics explain very little of the variation in performance for the math task in terms of goodness of fit measures such as R^2.

The results from the word task (Regressions 3 and 4) in Table 14 suggest that neither the tournament nor tournament size increase performance. There is a significant amount of learning that occurs with each treatment. GPA has a significant positive effect in terms of performance. This likely occurs because an individual's vocabulary expands with age and individuals with a higher GPA probably have richer vocabularies than individuals with lower GPAs. More of the variation in performance can be explained in regressions using the word task than the math task; this mainly stems from the inclusion of control variables for the random letters used for each task.

The math task results suggest that being in a tournament does increase performance, but the size of the tournament is irrelevant. One might expect that only high ability individuals would increase performance from the incentive effects from being in a tournament, but I find the opposite. In splitting the sample for high and low ability individuals, according to their performance in the first task and whether they are above or below the median, I find that the low ability individuals increase performance in response to being in a tournament (significant at 1%) in the math task. I find no significant effects from tournament size for high ability individuals in the math task.[7]

The competitiveness of the environment has a significant impact on performance only in math tasks and once an individual is participating in a tournament, then the number of competitors does not lead to further performance benefits. Competitive environments (tournament size) had no influence on the performance of individuals in the word task. Thus, depending on the type of task, competition between individuals may increase performance. Therefore, due to mixed results, one cannot conclude that tournament size increases performance or effort of agents.

Overall, I find that a more competitive work environment may not lead to performance increases as the incentive effects of competitiveness are not robust across different types of tasks. Another important result shown by these regressions is that there are no significant performance differences between males and females. In terms of performance effects, some learning occurs across the different treatments and only low ability individuals tend to increase performance when they are put in a tournament of any size– it is enough to be competing against someone.

As in other experiments that used similar tasks, there are no significant performance differences between males and females. This lack of difference suggests that both males and females should select into competitive environments in a a similar fashion, but I find that this is not necessarily the case. These selection differences are discussed in the following section.

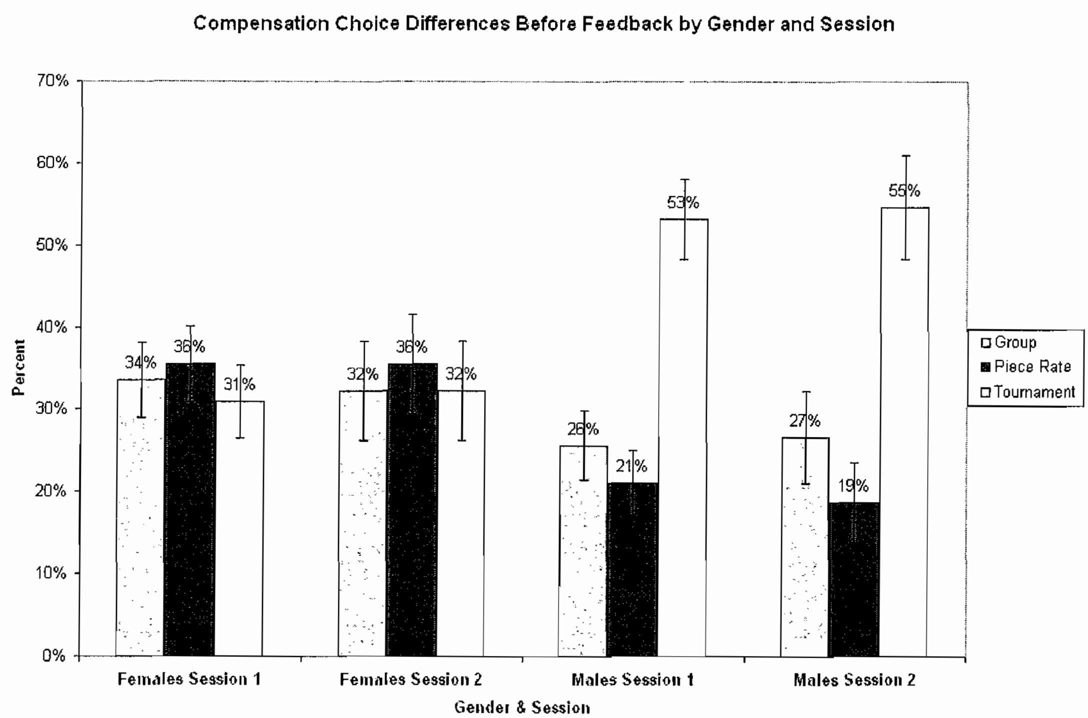
---

[7]The estimation results for high ability individuals are not shown here, but none of the competitive environment variables were significant in these estimations.

3.4.3 Gender Differences in Compensation Choice

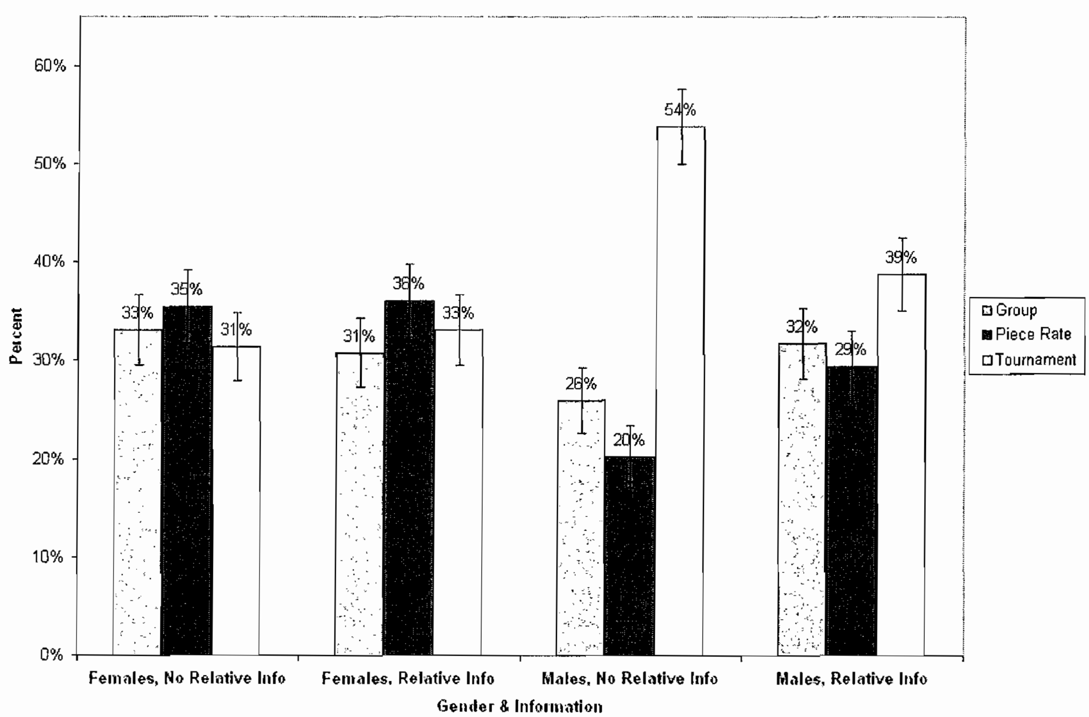
Previous experiments have found that females are less likely than males to enter tournament environments when given the option between a tournament and piece rate. To test whether these prior results can be replicated with this protocol, I first focus on choices made in Treatment 4. In this treatment, as in most previous studies, individuals were not given information about their relative performance against possible competitors from previous treatments.

FIGURE 6. Choices without Information by Gender and Session



Most individuals participated in two sessions. Figure 6 shows the histograms of the choices made by males and females in first and second sessions for Treatment 4. Both males and females behaved in a consistent manner in both sessions. The gender differences are large when no relative performance information was provided as over 53% of males in both sessions chose to compete in tournaments while only approximately only 30% of females select tournaments.

FIGURE 7. Choice Differences by Gender and Information



Error bars show standard errors.

Figure 7 shows the effect of relative performance feedback on competitive environment choices, which are aggregated by gender. In Treatment 4, subjects received no relative performance information; in Treatment 5, subjects received relative performance feedback prior to making their choices. Females tend to have an almost uniform distribution across choices as the proportions selecting group pay, piece rate, tournament are not significantly different from each other in both treatments. Females tend to select the piece rate more than the tournament or group pay, but this difference is not statistically significant. Females, when compared to males, systematically choose not to enter tournaments without relative performance information. Instead, females select both group pay and piece rate significantly more than males in Treatment 4. Thus, this treatment of no relative information replicates earlier findings while having an additional possible choice in the form of group pay.

The right side of Figure 7 shows that males select competitive environments quite differently when they have relative performance information about potential competitors. There is a significant increase in the proportion of males choosing the piece rate (5% level) and group pay (10% level), and a significant decrease in the proportion of males choosing the tournament (5% level). After receiving the relative performance information, the proportions of males and females have no statistically significant differences between them at the 5% level.[8] Relative performance information seems to remove any gender differences in competitive environment selections.

[8]At a 10% significance level there is a difference in the piece rate selection between males and females.

While there are no significant differences in performance between males and females, other factors such as age and GPA may play important roles in the output of individuals. If performance is affected by such factors, then individuals' choices to enter different competitive environments may also be affected by them. Therefore, it is important to test whether these gender differences in selection for Treatment 4 remain after controlling for these other potentially relevant factors.

Table 15 shows the results from ordered probit estimations from Treatment 4. Columns 1 to 3 use CompScale as the dependent variable where group pay is equal to -1, piece rate is equal to 0, and a selection of a tournament of any size is categorized as 1. As long as this ordinal ranking holds in terms of a dimension of competitiveness then this categorical variable can be properly used in an ordered probit estimation.

Table 15 shows that the results of Niederle and Vesterlund (2007) are replicated in the treatment without relative performance feedback: females choose to not enter competitive environments as frequently as males, even when controlling for individual confidence (Confidence(T1)) and relative rank of performance within the session (%-tile Rank(T1)) from the first treatment. This is shown by the negative significant coefficient for the female dummy variable (Female) in columns 1 through 3. Confidence is measured by the predicted performance an individual has at the end of Treatment 1 (prior to finding out their actual performance) divided by the individual's prediction of the average performance of all session participants. To control for performance, I use the relative rank from Treatment 1.[9] Both confidence

[9]In regressions, performance from the math task (word task) explains over 67% (57%) of the variation in the percentile rank in the math task (word task). The difference between the math and word task is likely attributed to the random assignment of letters for tasks.

TABLE 15. Ordered Probit: Choices with No Information

VARIABLES	(1) Pooled CompScale	(2) RE CompScale	(3) RE Risk CompScale
Female	-0.36	-0.40	-0.49
	(0.13)	(0.15)	(0.19)
	(***)	(***)	(**)
Confidence(T1)	0.86	0.98	0.99
	(0.25)	(0.29)	(0.34)
	(***)	(***)	(***)
Improve (T2)	0.61	0.72	0.73
	(0.20)	(0.23)	(0.32)
	(***)	(***)	(**)
%-tile Rank (T1)	1.05	1.08	0.85
	(0.23)	(0.26)	(0.32)
	(***)	(***)	(***)
Risk Controls	No	No	Yes
Characteristic Controls	Yes	Yes	Yes
Observations	343	343	224
ll	-336.6	-335.6	-212.3
chi2	66.91	61.00	48.81

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

and the percentile rank from treatment one are positively correlated with the selection of more competitive environments. Improvement in task performance between the first and second task (Improve(T2)) also has a significant positive effect on choosing a more competitive environment in Treatment 4.

Table 15 shows that females are less likely to enter into competitive environments even when controlling for confidence, performance, improvement and some individual specific characteristics such as the number of years of college education, usage of psychological medication, GPA and age. The results are similar when using a random effects ordered probit as shown in column 2. The possibility exists that gender differences in the self selection of competitive environments are in part attributed to risk aversion. Column 3 includes measures of risk aversion for individuals that participated in a task similar to the one used by Holt and Laury (2003). I find that risk aversion is not significantly correlated with competitive environment choice.

3.4.4 Do Differences in Choices Exist with Information?

With no information concerning the quality of potential competitors, females choose to enter less competitive environments than males, irrespective of confidence, performance, improvement or risk aversion. These gender differences do not seem to be related to risk or overconfidence suggesting that these differences are related to preference differences for competition. If a preference difference for competitive environments exists then providing information about the quality of possible competitors might reduce mistakes in selections of environments that are too competitive or not competitive enough, but this information should not reduce the gender differences in choices. If females dislike competitive environments more than males then there should still be a negative gender effect for females once both males and females are informed of the quality of possible competitors in Treatment 5.

In Treatment 5, before choosing the competitive environment, individuals are shown a list of the performance of all the participants in the session from Treatment 1. This information provides a signal of the quality of competitors that participants may compete against if they choose to enter a tournament setting. Or, alternatively, a signal of the quality of person they may be paired with if they choose group pay.

Table 16 shows the results from two different types of ordered probit estimations for Treatment 5 where individuals receive relative performance information prior to making choices about compensation. The table shows that females do not make significantly different choices than males for a number of different specifications. Columns 1 and 2 provide estimates using CompScale as the dependent variable, where there are only three dimensions, group pay, piece rate, and tournament. Using this dependent variable, I find there are no differences between males and females in the selection of competitive environments. The selection of competitive environments is mainly dependent on the relative performance information provided prior to making the decision and an individual's improvement from Treatment 1 to Treatment 2. Risk aversion control variables are not significantly related to whether an individual chooses a tournament, piece rate, or group pay compensation scheme. Though, the sample of individuals that participated in the risk aversion task, whether due to risk aversion controls or sample differences, seem to not be affected by improvement between tasks to the same degree (as magnitude and significance are both lower) as the rest of the sample. The one variable that consistently matters for individuals in this treatment is their percentile rank in Treatment 1.

There are no significant differences between females' and males' selections of competitive environments in Treatment 5 when individuals were fully informed of their relative performance compared to potential competitors. The lack of differences between males and females in this treatment raises questions about the existence and

TABLE 16. Ordered Probit: Choices with Information

VARIABLES	(1) Pooled CompScale	(2) RE CompScale	(3) RE Risk CompScale
Female	0.00	-0.02	0.13
	(0.13)	(0.18)	(0.21)
	()	()	()
Confidence (T1)	0.34	0.44	0.65
	(0.24)	(0.30)	(0.35)
	()	()	(*)
Improve (T2)	0.81	1.01	0.65
	(0.20)	(0.26)	(0.32)
	(***)	(***)	(**)
%-tile Rank (T1)	2.17	2.59	2.31
	(0.25)	(0.34)	(0.37)
	(***)	(***)	(***)
Risk Controls	No	No	Yes
Characteristic Controls	Yes	Yes	Yes
Observations	343	343	224
ll	-320.6	-316.7	-194.5
chi2	110.9	98.51	79.67

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

sources of gender differences for competitive environments. This effect of information suggests that, relative to males, females may be less likely to compete when they lack knowledge about the quality of the competition. In the section below, I 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.

3.4.4.1 The Cost of Gender Differences Choices

To compare the costs of gender differences for competitive environments, I approach the problem as having each individual maximize expected value. This approach does not take into account risk aversion, but focuses on the dollar costs of suboptimal choices. In the case where there is no relative performance information, if females are making compensation choices by avoiding tournaments for promotion then these selections may be costly. The results have shown that with poor relative performance information, females select not to enter tournaments, while males overwhelmingly select to enter tournaments. Both types of selections may be suboptimal as males may select competitive environments too much and females may be selecting tournaments too little.

3.4.4.2 Expected Value of Choices

To maximize expected value from choosing to compete in a tournament, or to accept the piece rate, or to select the group pay scheme, an individual cares about his probability of winning a tournament and his expected output. Expected values are based on the performance from Treatment 1 and the improvement of the individual from repeating the task and being in a tournament in Treatment 2.[10] Assume an exogenous probability of winning a two person tournament in the form the percentile rank from Treatment 1 for the individual. If the individual is the best in the session then he receives a probability of winning a tournament of 0.99 and if he is the worst

[10]For simplicity, the incentive effects of effort from entering a tournament are ignored in calculating these expected values.

performing individual in the group his probability of winning is set to 0.01. Let this rank be equal to p. An individual's probability of winning an n person sized tournament is then p^{n-1}. Let b be the base piece rate and y be the expected output of the individual, then the expected value of an n person tournament is:

$$EV_n = p^{n-1}\, y\, b\, n \tag{3.4.1}$$

This expected value form includes the piece rate which is the equivalent of a tournament size of 1. For group pay selection, assume that an individual accurately predicts the mean performance of the group. If an individual expects his output to be lower than the group average and if his probability is less than $\frac{1}{2}$ then he should choose the group pay compensation scheme. Otherwise the rank ordering for the different sized tournaments in terms of expected values is as follows:

$$\left\{ \begin{array}{ll} EV_6 \geq EV_4 & \text{if } y > \bar{y} \text{ and } p \geq \sqrt{\frac{2}{3}} \\ EV_4 \geq EV_6 & \text{if } y > \bar{y} \text{ and } p \leq \sqrt{\frac{2}{3}} \\ EV_2 \geq EV_4 & \text{if } y > \bar{y} \text{ and } p \leq \sqrt{\frac{1}{2}} \\ EV_1 \geq EV_2 & \text{if } y > \bar{y} \text{ and } p \leq \frac{1}{2} \\ EV_{Grp} \geq EV_n & \text{if } y \leq \bar{y},\ \forall n \in \{1, 2, 4, 6\} \end{array} \right\} \tag{3.4.2}$$

3.4.5 Costly Choices

In this experiment, high ability individuals and low ability individuals make choices about the form of compensation they will receive. To maximize the returns from effort, a high ability individual should choose a more competitive environment while a low ability individual should choose a less competitive environment. An individual should choose a more competitive environment the higher that individual's relative ability is compared to the group of possible competitors. This choice should occur because a higher ability individual is more likely to win a tournament. Furthermore, in this experiment there is an even greater incentive for choosing a tournament for high ability individuals because individuals that choose to compete in a tournament may be randomly chosen to compete against individuals of lower ability who did not choose to enter the tournament. Thus, this framework is structured towards having higher ability individuals choose a more competitive tournament and lower ability individuals to choose a piece rate or group pay form of compensation.

Table 17 shows the average expected value losses for suboptimal selections by males and females in Treatment 4. Each row represents an actual decision that was made by individuals, while each column represents the optimal decision that should have been made. For example, the first row and column section (Grp Loss, Grp) cell shows the average loss for individuals that selected the group pay and should have selected group pay. In this case, the optimal choice was group pay and the average loss from optimally choosing group pay is zero. In contrast, under column 6 (for the 6 person sized tournament), the average expected value loss was 19.82 for males whose optimal choice would have been to choose a tournament of six, but who chose group pay instead.

In Treatment 4, the average expected value loss of selection mistakes was 4.91 for males and 6.78 for females.[11] These loss differences are mostly driven by high

[11] This difference is significant at a 10% level using a 1-tailed test.

TABLE 17. Treatment 4 Selection Loss

Males Treatment 4 Selection Loss						
		Optimal Choice				
Chosen	Grp	1	2	4	6	Avg Loss
Grp Loss	0	0.89	2.45	5.90	19.82	1.74
1 Loss	0.96	0	1.51	7.69	11.71	2.71
2 Loss	5.84	2.65	0	0.40	37.93	11.75
4 Loss	7.28	5.27	1.58	0	15.49	8.05
6 Loss	6.63	7.51	5.74	0.53	0	3.37
Avg Loss	2.42	2.97	2.31	3.29	12.60	4.91
Females Treatment 4 Selection Loss						
		Optimal Choice				
Chosen	Grp	1	2	4	6	Avg Loss
Grp Loss	0	0.48	2.16	5.00	42.61	5.18
1 Loss	0.95	0	1.48	9.52	39.28	9.98
2 Loss	4.57	2.26	0	1.81	.	3.19
4 Loss	9.27	6.51	3.91	0	7.31	5.51
6 Loss	7.08	9.04	6.97	.	0	5.43
Avg Loss	1.58	2.28	2.91	6.80	27.27	6.78

ability females and to a lesser extent by low ability males. It becomes apparent from examining column 6 for males and females, that high ability females (those who should select a tournament size of 6) are making the costly choice of selecting group pay or piece rate. Females lose 27.30 in expected value compared to 12.60 for males. In contrast, low ability males make more costly decisions than low ability females as the females that should select group pay lose 1.58 while the low ability males lose 2.42 on average. I find that both high ability females and males are not entering competitive environments enough. But these high ability females overwhelming select the noncompetitive environments of piece rate and group pay, which are very costly decisions. In contrast, too many low ability males are entering competitive environments.

Table 18 shows that the relative performance feedback decreases the average expected value losses for both males and females. In Treatment 5, there are no longer significant differences in terms of expected value losses between males (3.95) and females (4.80) from suboptimal choices. The decreases in expected value losses are greatest for high ability females; their average expected value loss improved to 18.70 in Treatment 5 from 27.27 in Treatment 4. High ability males also made more value maximizing decisions as the average expected value loss for males that should have selected a tournament size of six improved to 10.98 in Treatment 5 from 12.60 in Treatment 4. Thus, high ability females are still not entering competitive environments enough, but they are making better value maximizing decisions when they receive relative performance information by entering more competitive environments.

TABLE 18. Treatment 5 Selection Loss

Males Treatment 5 Selection Loss						
			Optimal Choice			
Chosen	Grp	1	2	4	6	Avg Loss
Grp Loss	0	0.49	2.27	7.52	7.53	0.84
1 Loss	1.28	0	1.63	6.02	21.29	4.54
2 Loss	6.39	2.16	0	.	32.88	8.61
4 Loss	5.76	5.52	2.93	0	10.20	7.59
6 Loss	5.16	9.59	6.64	0.77	0	2.55
Avg Loss	1.39	1.49	2.02	4.79	10.98	3.95
Females Treatment 5 Selection Loss						
			Optimal Choice			
Chosen	Grp	1	2	4	6	Avg Loss
Grp Loss	0	0.46	1.80	4.65	.	0.46
1 Loss	0.85	0	1.74	8.18	32.34	7.77
2 Loss	4.24	2.65	0	5.57	18.51	6.99
4 Loss	7.32	6.35	3.39	0	11.39	6.27
6 Loss	.	6.68	4.22	3.33	0	2.64
Avg Loss	0.88	1.88	2.21	5.93	18.70	4.80

Low ability females and males tend to behave similarly after receiving the relative performance information as more of them optimally select group pay. The low ability females in the Treatment 4 lose (1.58) significantly less expected value than males (2.42).[12] Once these low ability individuals receive relative performance information, they make decisions that increase expected value, but a gender difference in terms of costs still remains as the expected value losses are 0.88 for low ability females and 1.39 for low ability males.[13] Therefore, the costs of suboptimal selections by both genders improve substantially with relative performance information.

Figure 8 provides the distributions of choices made by both high ability females and high ability males. High ability is defined as an individual that should enter a four person tournament and low ability is defined as someone who's performance is below the median performance level of their respective session. Figure 8 shows that the relative performance information leads to an increase in the proportion of high ability females entering tournaments. Over 50% of high ability females enter tournaments as compared to 31% without information. With information, fewer high ability males enter tournaments (12% less), but this change in tournament selection is not statistically significant (at 5% level). Relative performance information leads to more high ability females entering tournaments.

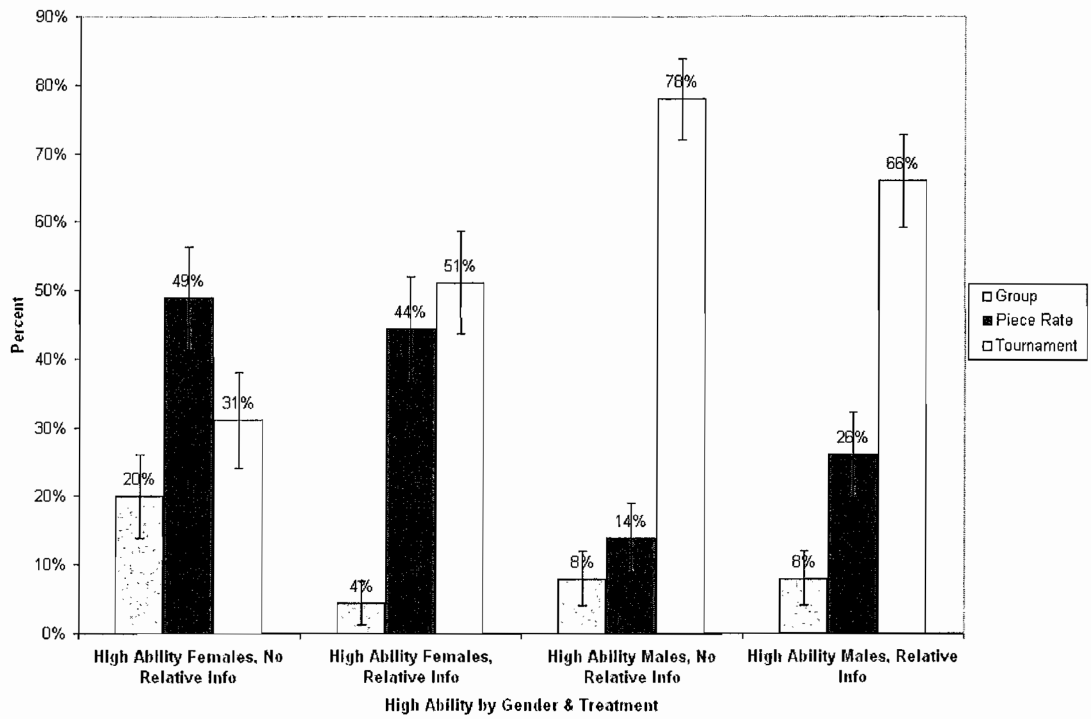
Figure 9 shows that low ability males enter into more competitive environments than females when they have no relative performance information. Once these low ability males can make decisions using the relative performance information, the

---

[12] 1-tailed p=0.063

[13] 1-tailed p=0.097

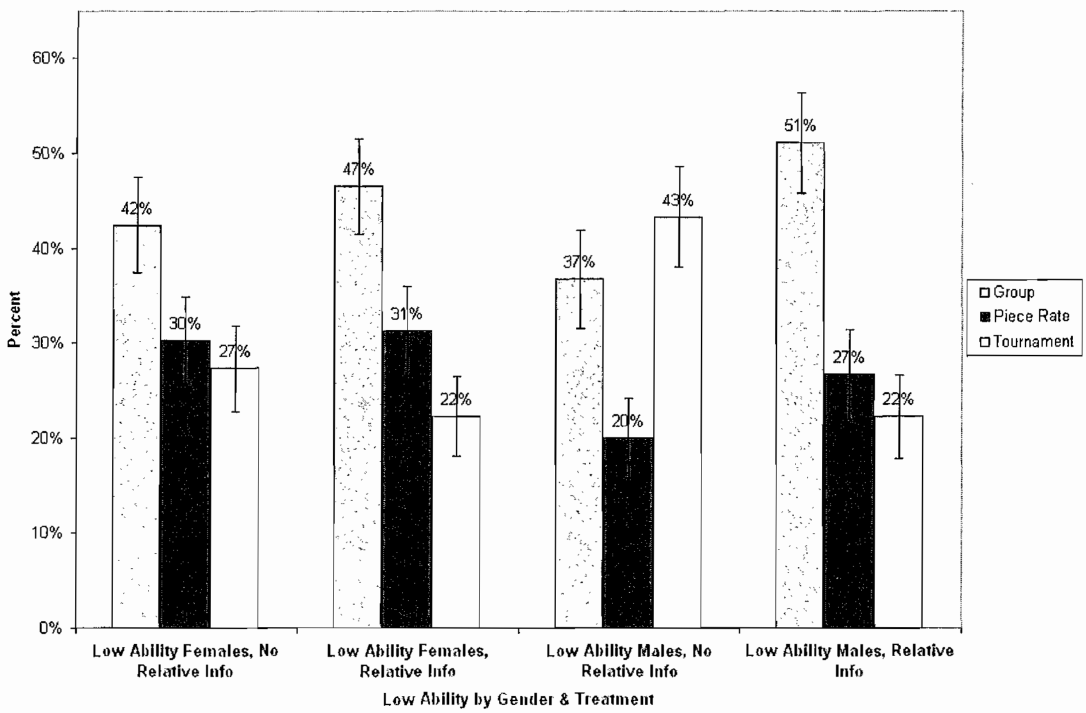
FIGURE 8. Information Effects for High Ability



Error bars show standard errors.

proportion in tournaments drops from 43% to 22%. The information leads to 51% of these low ability males to select group pay, a significant increase from the 37% observed without the relative performance information. In considering low ability individuals, relative performance information mostly leads to more low ability males moving out of tournaments and towards the group pay form of compensation while low ability females exhibit little change.

FIGURE 9. Information Effects for Low Ability



Error bars show standard errors.

High ability females select more competitive environments when given relative performance feedback. Fewer high ability males select the tournament with information, but on average the expected value loss is lower for high ability males once they receive relative performance information. Females could still choose more competitive environments, but the information has at least moved most high ability females away from selecting group pay. Low ability males also exhibit large changes in behavior after receiving the relative performance information as many of them switch away from choosing tournaments towards the optimal choice of group pay. Thus, the relative performance information leads to more efficient sorting by both males and females.

Without relative performance information, males and females select competitive environments differently, but there are other possible biological reasons for differences in choices within the female population. The following section will examine whether the hormone phases of the menstrual cycle for women are correlated with competitive environment choices. If there is a systematic relationship between phases, then this may provide a further explanation for the drastic gender differences shown by females choices in the treatment without relative performance information in this experiment.

3.4.6 Hormonal Effects for Competitive Environments

Normal cycling women experience large changes in hormones across the menstrual cycle, and hormonal contraceptives alter the form of these fluctuations for females. Both groups of females experience a low hormone phase, this is during menses for normal cycling females and during the placebo phase (or non-use) for females using a hormonal contraceptive. I exploit this change in hormones in both groups of females to test whether hormonal fluctuations are related to competitive environment choices.

Females using hormonal contraceptives experience more predictable hormonal fluctuations than females who do not use hormonal contraceptives. Normal cycling females, those not using hormonal contraceptives, tend to experience more noise in the timing of their menstrual cycle period. Females using a hormonal contraceptive do not ovulate and during the off-week (or placebo phase) experience withdrawal bleeding caused by low-hormone levels. Thus, the low-hormone phase for both hormonal contraceptive users and normal cycling females occurs during the week of their menstrual period or withdrawal bleeding.

To identify where each participating female was in her menstrual cycle, screening and experiment session exit surveys were used. The screening survey was used to schedule all female participants during both a low-hormone phase and high-hormone phase. The high-hormone phase corresponds to the mid-luteal peak for normal cycling females (Figure 5) or three days into contraceptive use for female using a hormonal contraceptive. These phases were verified in the experiment exit survey that was filled out by each participant as experiment payoffs were prepared. Accurate scheduling of females according to their menstrual cycles was an important aspect of the study. Thus, it was important to know the variability of participants' menstrual cycles, forms of contraception used and start days of hormonal contraceptive regimens.

TABLE 19. 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

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

Table 19 provides a summary of the answers given by potential female participants in the screening survey concerning the regularity of their menstrual periods. Some females experience a large degree of variability in their menstrual cycles. Of the females who completed the screening survey, almost 15% missed a menstrual period during the previous 3 months. Consequently, scheduling around predicted phases would not be accurate enough to identify hormonal phases. For this reason, an exit survey was also used at the end of each experiment session to help with identification of hormonal phases.[14]

TABLE 20. Female Birth Control Use

Form of Birth Control	Percent	Count
Abstinence	11.3%	44
Pill	48.7%	189
Condoms	35.3%	137
Depo-provera	0.8%	3
IUD*	2.3%	9
Patch	0.0%	0
Timing	1.8%	7
Tubal Ligation	0.3%	1
Vaginal Ring	5.9%	23
Vasectomy	0.3%	1
Other	2.1%	8
None	19.3%	75
Total		497

Multiple responses were allowed.

*Intrauterine device

[14]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 phase in the hormonal cycle.

As mentioned in Section 3.2.2, hormonal contraceptives provide women with a choice about whether to menstruate. Many females already use a hormonal contraceptive to keep a regular and predictable menstrual period.[15] To understand how widespread hormonal contraceptive use is, at least among college students, consider the data in Table 20 obtained from females in the screening survey concerning the forms of birth control used. Over 54% of females use some form of hormonal contraceptive. Thus, over half the potential female participants could avoid the low-hormone phase by not taking an off-week or placebo week as part of their hormonal contraceptive regimen.

To help identify hormonal phases for females using a hormonal contraceptive, the start day for hormonal contraceptive use was asked of all potential female participants. Table 21 shows that females who completed the screening survey overwhelmingly favor a Sunday start day for their hormonal contraceptive regimen. This allows for easier predictability of low and high phases for these females as hormonal fluctuations are exogenously determined by hormonal contraceptive use.

Figure 10 shows the distribution of choices of females across the three possible competitive environments of group pay, piece rate and tournaments for Treatment 4, where participants had no relative performance information prior to making their choices. The histogram includes all females that were able to attend two sessions and could be identified as being in a low or non-low-hormone phase. Many of these females

[15]With hormonal contraception, the bleeding females experience is withdrawal bleeding and not a menstrual period.

TABLE 21. Hormonal Contraceptive Start Day

Start day	Percent	Count
Sunday	51.7%	106
Monday	11.2%	23
Tuesday	8.3%	17
Wednesday	9.3%	19
Thursday	8.8%	18
Friday	7.3%	15
Saturday	3.4%	7
Total		205

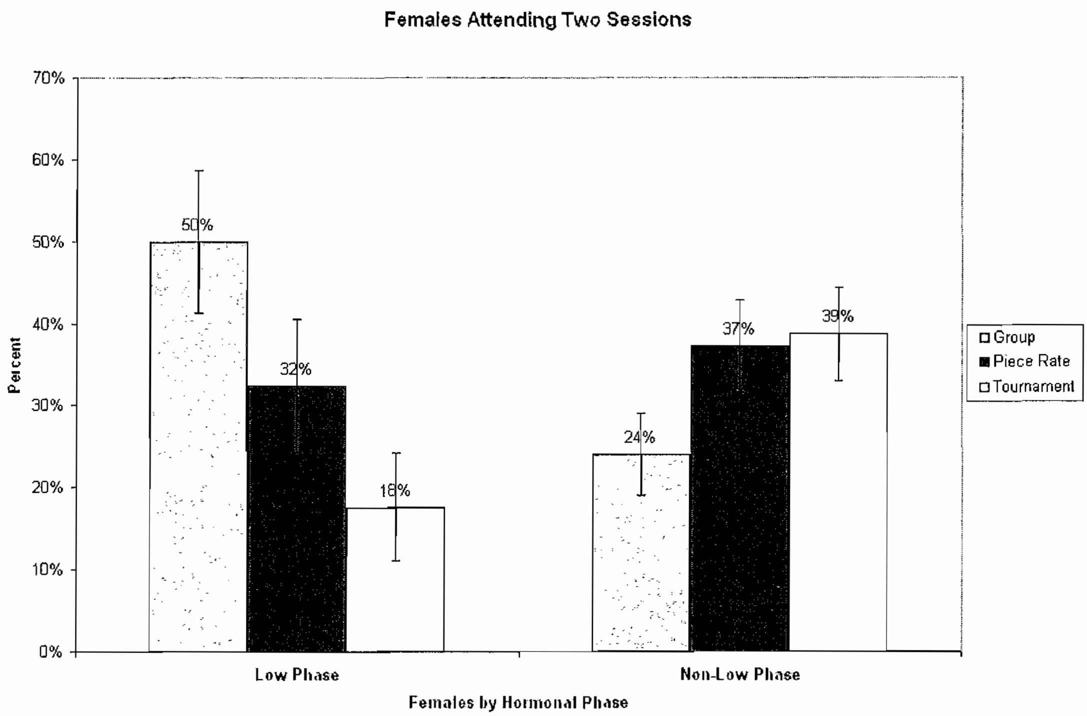
were scheduled for both a predicted low and contrasting high hormonal phase.[16] The figure suggests that low phase females are more likely to choose group pay and less likely to choose tournaments in comparison to females in a non-low-hormone phase when they have not received any relative performance feedback.

3.4.7 Performance differences by hormonal phase.

These differences in competitive environment choices between females in the low-hormone phase and females in the non-low-hormone phase may be related to decreased expected performance across the menstrual cycle or from different preferences for competition in the math and word task. I 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. To examine performance differences across the menstrual cycle, I consider the word and math tasks separately and estimate effects using linear specifications similar to the ones used to examine exogenous performance effects of

[16]Due to the error in predicting the low phase some females were in a non-low phase for both word and math tasks.

FIGURE 10. Compensation Choice & Hormonal Phase for Females



Error bars show standard errors.

tournaments. Table 22 provides the random effects OLS estimates for a number of factors that may explain performance differences in both word and math tasks for all treatments where the participants could not choose their competitive environments. The estimations for math tasks are in columns 1 to 3 and the estimations for word tasks are in columns 4 to 6.

There seems to be no correlation with the low-hormone phase and performance in the word task (columns 4 to 6) for females. The Low Phase coefficient is insignificant for all the different samples in the word task. Focusing on performance in the math task (columns 1 to 3), there seems to be no effect from the low phase in the sample of both females and males (column 1), and only females (column 2). There seems to be some effect for low phase females when including controls for risk aversion for the portion of the particpants for which such measures were available.

On average, performance in the math task decreases by about 2.2 correct answers for low phase females when controlling for risk aversion, which is a 20% decrease for the average female. Though the low phase effect is only significant at the 10% level when taking into account individuals for which measures of risk aversion can be used as controls, it still suggests that anticipated performance differences may play some role in differences in selection exhibited by females in the low phase. Therefore, if performance or confidence differences are driving selection differences, then controlling for performance and confidence in a discrete choice model should help isolate the effect of the low phase on selection choices.

TABLE 22. Hormonal Effects for Performance (t<4)

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	Female	Female	All	Female	Female
Task	Math	Math	Math	Word	Word	Word
VARIABLES	RE	RE	RE Risk	RE	RE	RE Risk
Task Order	0.87	0.94	1.01	3.46	3.92	3.09
	(0.09)	(0.13)	(0.18)	(0.39)	(0.57)	(0.70)
	(***)	(***)	(***)	(***)	(***)	(***)
Tourney Size	0.14	0.12	0.09	0.05	-0.03	-0.20
	(0.04)	(0.06)	(0.08)	(0.16)	(0.23)	(0.29)
	(***)	(**)	()	()	()	()
Low Phase	-1.07	-1.09	-2.21	2.22	2.09	4.57
	(0.98)	(0.71)	(0.89)	(3.00)	(2.98)	(3.58)
	()	()	(**)	()	()	()
Female	-0.32			1.72		
	(0.67)			(2.11)		
	()			()		
Years PS	-0.41	-0.70	-1.19	0.31	-0.87	-3.41
	(0.30)	(0.40)	(0.52)	(0.87)	(1.39)	(2.10)
	()	(*)	(**)	()	()	()
Age	0.38	0.55	0.89	0.31	1.48	3.52
	(0.15)	(0.23)	(0.27)	(0.49)	(0.90)	(1.22)
	(**)	(**)	(***)	()	()	(***)
GPA	-0.12	-0.43	-0.46	5.73	7.56	6.22
	(0.65)	(0.78)	(1.03)	(2.00)	(3.19)	(4.13)
	()	()	()	(***)	(**)	()
Psych meds	-1.48	-0.88	-1.17	3.37	4.80	3.56
	(1.10)	(1.19)	(1.65)	(3.23)	(4.95)	(6.78)
	()	()	()	()	()	()
Constant	3.26	0.78	-3.87	6.92	-18.13	-43.30
	(3.47)	(4.18)	(5.28)	(11.70)	(19.47)	(26.05)
	()	()	()	()	()	(*)
Letter Controls	No	No	No	Yes	Yes	Yes
Risk Controls	No	No	Yes	No	No	Yes
Observations	510	237	147	471	225	147
Number of id	170	79	49	157	75	49
R-sq o	0.0909	0.168	0.339	0.369	0.410	0.496
chi2	100.5	62.28	54.45	563.4	289.1	209.3

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

3.4.7.1 Differences Between Word and Math Tasks for Hormone Phases.

The distribution of choices made by females across the menstrual phase suggest that hormonal phases may play a role in competitive choice differences within the female population. Females in a specific hormonal phase may also experience greater aversion to certain types of tasks; therefore, I examine whether females in the low-hormone phase are more likely to choose group pay and less likely to choose a tournament for both math and word tasks.

FIGURE 11. Choices by Phase for Females that Attended Both Sessions.

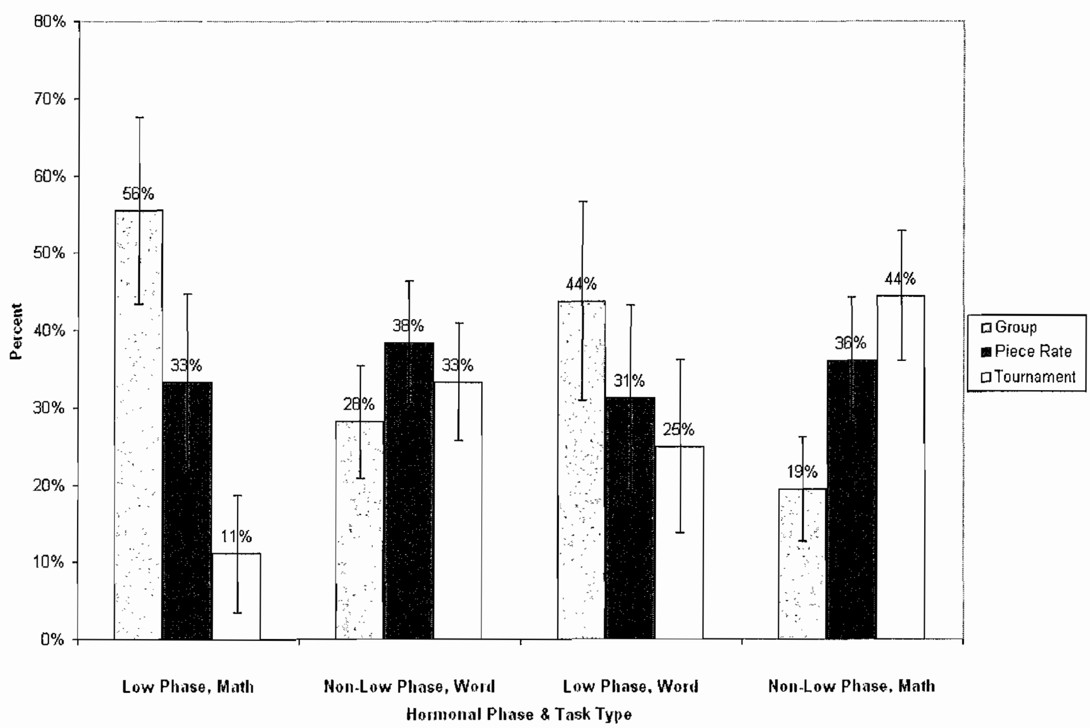


Figure 11 provides histograms of females' choices by hormonal phase and task type. The general pattern of low phase females is the same for females that participate in word and math tasks during the low phase. Low phase females are more likely to choose group pay and less likely to choose tournaments in both math and word tasks. Non-low phase females, exhibit slightly different choices between math and word tasks as they overwhelmingly select tournament in the math task, while in the word task, tournament is still selected more than group pay, but piece rate is selected the most. Non-low phase females consistently select group pay the least in both math and word tasks, which is a direct contrast to females in the low-hormone phase. Subjects alternated the type of task used in sessions; females that participated in a math or word task during the low phase would then participate in the other type of when not in the low phase, and vice-versa. The differences in selections between hormonal phases suggest that these phases may be correlated with behavioral differences for females.

3.4.7.2 Do Low Phase Females Make Different Choices?

Table 23 provides the results from ordered probit models for Treatment 4 choices using the CompScale variable, which is an ordered categorical variable consisting of the following categories: group pay, piece rate, and tournament.[17] I find that females in the low phase choose to enter less competitive environments than females in a non-

[17]Estimations using the tournament size ordered categorical variable that takes into account further dimensions of competitiveness were not robust as significance for the low phase fluctuates depending on the specification. Too few low phase females chose tournaments to allow for disaggregation by tournament size.

low phase. Without relative performance information, much of the difference between genders seems to be driven by low phase females, who shy away from the competitive environments of tournaments and choose the least competitive setting possible in the form of group pay.

TABLE 23. Ordered Probit: Hormone Effects for No Relative Information (Treatment 4)

	(1)	(2)	(3)	(4)
Sample	All	All	Females Only	Risk
VARIABLES	Pooled	RE	RE	RE
Female	-0.26	-0.29		-0.26
	(0.14)	(0.16)		(0.21)
	(*)	(*)		()
Low Phase	-0.44	-0.46	-0.53	-0.76
	(0.21)	(0.22)	(0.26)	(0.27)
	(**)	(**)	(**)	(***)
Confidence (T1)	0.81	0.91	1.08	0.90
	(0.26)	(0.30)	(0.54)	(0.35)
	(***)	(***)	(**)	(**)
Improve (T2)	0.60	0.69	0.79	0.72
	(0.20)	(0.23)	(0.38)	(0.32)
	(***)	(***)	(**)	(**)
%-tile Rank (T1)	0.97	0.99	0.52	0.72
	(0.23)	(0.26)	(0.43)	(0.32)
	(***)	(***)	()	(**)
Risk Controls	No	No	No	Yes
Characteristic Controls	Yes	Yes	Yes	Yes
Observations	328	328	155	211
ll	-322.3	-321.7	-156.0	-197.4
chi2	64.32	58.60	19.76	51.31
r2_p	0.0907			

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The total number of of low phase females that can be used for data analysis is 45.

Table 24 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. The first column provides pooled cross-sectional results including all subjects, the second to fourth columns provide estimates from using a random effects ordered probit. The second column includes all males and females, the third column consists of a female only sample and the fourth column takes into account only males and females for which risk aversion measures were available. The results in Table 24 show 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. Once all participants are informed of the quality of possible competitors then differences between genders and across the menstrual cycle become insignificant.

There is a cost associated with high ability individuals not entering competitive settings and with low ability individuals not choosing less competitive settings. Low-hormone phase females make more costly decisions compared to non-low phase females and males when using expected value loss calculations. The expected value loss compares the expected value of optimal competitive environment selection with the actual selection expected value.[18] The average expected value losses for males, low phase and non-low phase females in Treatment 4 are shown in Table 25. Treatment 4 is used because in Treatment 5 selection differences are insignificant. In Treatment 4, low phase females sacrifice the greatest amount of expected value, \$8.50 per decision. The expected value losses for males and non-low phase females are respectively \$4.91

[18]These expected values are calculated in the same way as discussed in the previous section.

and \$6.52. The value difference between low and non-low phase females is not statistically significant and the difference between males and non-low-phase females is not statistically significant either. But low phase females make more costly choices than males (5% significance).

These results imply that hormones may matter in the selection of competitive environments, but only if the strength of the competition or the probability of winning is relatively unknown. If there is little information, 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.[19]

3.4.7.3 Absenteeism, Cancellations and Tardiness

This study was structured to have a large number of low-hormone phase females to compare with the rest of the sample. Due to the noisiness of the menstrual cycle, it is difficult to predict the low hormonal phase for females. Using the screening survey, female subjects were scheduled for their sessions according to their predicted cycle day. This was calculated using self reported data about the start of subjects' previous menstrual periods. Whenever possible these data were combined with self reported data concerning females' hormonal contraceptive regimens. Once the cycle day could be predicted, then a set of possible session days were provided to potential participants and they chose and confirmed these days with a research assistant. For

[19]These insignificant differences are not shown here.

individuals to be scheduled, they had to confirm that they would attend a specific session.

Absenteeism, cancellations and tardiness are frequent in experiments. In this study, because of the screening survey and experiment exit surveys, something can be learned about the individuals who do not show up for their scheduled session or cancel at the last moment or are tardy. Ichino and Moretti (2009) found that female worker absenteeism was highly correlated with the female menstrual cycle at a regular job. Due to systematically greater incidences of absences and cancellations in predicted low phase females in this experimental setting, I find results that support their findings.

Table 26 shows that females are significantly less likely to show up as scheduled when compared to males. The table shows the proportion of participants that attended the experiment sessions as scheduled, meaning they were present and punctual. 79% of the males showed up as scheduled. Based on predicted phases, only 62% of low phase females low phase females attended as scheduled while non-low phase females attended 72% of the time. These differences between attendance rates of low phase and non-low phase females were significant (5% level). Furthermore, non-low phase females were significantly (10% level) less likely to attend compared to males. Thus, there is both a gender difference in attendance as well as a difference in attendance related to the phase of the female menstrual cycle.[20]

[20] Attendance rates suggest that 62 low phase females should be identified in the experiment, but due to error in predicting the female menstrual cycle only 45 session participants can be classified in the low phase.

These attendance results suggest there may be a systematic bias in the hormonal phase of females who show up to experiment sessions. Additionally, if females who do not show up to sessions in this study are the ones who have worse symptoms during the low phase, or are more likely to behave differently, then a selection bias may add a downward bias to the hormonal effects found in this study. Since this study was partly focused on hormonal fluctuations, attempts were made to incentivize more low phase females to attend once this bias was found to exist. Part way through the study, due to low attendance from scheduled predicted low phase females, the participation payment of $5 was raised to $10 for low phase females to induce greater attendance of these females.[21] Even with this increase in participation payments for low phase females, a significant and systematic difference exists between individuals who attended sessions as scheduled and confirmed.

3.5 Discussion

In this study, I show that gender differences in compensation choices that have been found 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. This consistency is further shown by the similar choices made by individuals who participated in a second iteration of the experiment. I also find that females' choices vary across the menstrual cycle: in a low-hormone phase, females are less likely to enter tournaments than during a non-low phase. However, once relative performance

[21]These females were not informed as to the reason they were receiving a higher participation payment.

information is provided to subjects, all these differences for competitive environment selections become smaller and statistically insignificant.

Previous studies have shown that the gender differences in compensation choices persist even after controlling for confidence and have concluded that an underlying gender difference in the taste for competition must drive the result. But if the choice differences originate from a difference in preferences for competitive environments then the selection differences should remain even after relative performance feedback is provided. Another possible explanation would be a gender difference in risk aversion. However, I control for risk aversion and beliefs about relative ability. With these controls, a gender difference still remains in the uninformed treatment, but is removed once subjects receive relative performance feedback. These results suggest that a lack of knowledge about the distribution of the quality of potential competitors causes males and females to make very different choices. Kagel and Roth (1995) define ambiguity as *"known to be missing information."* Thus, an explanation that would be consistent with these results is that males and females make compensation decisions differently when the level of ambiguity is high; once ambiguity is removed they behave similarilly.

I have also shown that hormonal fluctuations affect female entry into competitive environments when there is no relative performance information. This result can be particularly valuable for females to consider for important decisions such as choosing graduate programs or career choices. Females can adapt behaviors or make choices to mitigate or remove the effects of these hormonal fluctuations. Females are able to control these fluctuations by changing their hormonal contraceptive regimen such that they would not experience a low-hormone phase. Females may also refrain from making competitive environment choices during the low-hormone phase. Such behaviors would only have value if there is poor information about the quality of individuals that are part of the competitor pool.

Affirmative action policies are often used to encourage more females to participate in competitive environments in the workplace and in educational institutions. These policies typically involve increasing efforts to recruit females or changing the acceptance or promotion process to favor females. The results from this study suggest that providing better information to applicants and labor force participants about their relative abilities may also reduce the gender differences. Many work and educational environments where affirmative action policies are implemented involve a great deal of ambiguity concerning relative abilities of individuals in the employee or student pool. If feedback eliminates the gender differences in compensation choices then it may be possible to change gender specific participation rates in competitive environments without necessarily altering the structure of the recruitment, acceptance or promotion process.

3.6 Concluding Remarks

This study has called into question whether a basic female distaste for competition is driving selection differences. Instead, these results suggest that males and females may process information differently, and make decisions differently, when ambiguity exists. These selection differences in the absence of information are heightened for females in the low-hormone phase of their cycle. Thus, gender differences for competition are not just dependent on the informational environment, but also on the hormonal cycle of females.

Although this study has shed some light on the possible causes of gender differences with respect to choices for competitive environments, there is still a long way to go before these differences between males and females are thoroughly understood. Relative performance information, or the lack there of, seems to play a vital role in these selection differences. One cannot say that the choice differences between females and males are caused by hormonal differences, but due to the correlations of choices with low and high hormone phases, this study does suggest that hormones may play a role in compensation choices for females. To accurately identify whether hormones are a determining factor for these choice differences, further study will require the measurement of hormone levels within subjects.

TABLE 24. Ordered Probit: Hormone Effects for Relative Information (Treatment 5)

	(1)	(2)	(3)	(4)
Sample	All	All	Females Only	Risk
VARIABLES	Pooled	RE	RE	RE
Female	-0.07	-0.13		0.10
	(0.15)	(0.20)		(0.24)
	()	()		()
Low Phase	0.03	0.13	0.12	-0.12
	(0.21)	(0.25)	(0.27)	(0.29)
	()	()	()	()
Confidence (T1)	0.23	0.32	0.21	0.54
	(0.25)	(0.31)	(0.53)	(0.36)
	()	()	()	()
Improve (T2)	0.76	0.92	1.06	0.49
	(0.20)	(0.26)	(0.40)	(0.33)
	(***)	(***)	(***)	()
%-tile Rank (T1)	2.18	2.61	2.63	2.33
	(0.25)	(0.35)	(0.55)	(0.38)
	(***)	(***)	(***)	(***)
Risk Controls	No	No	No	Yes
Characteristic Controls	Yes	Yes	Yes	Yes
Observations	328	328	155	211
ll	-307.8	-303.9	-143.3	-183.4
chi2	104.7	93.82	45.89	75.79
r2_p	0.145			

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

TABLE 25. Expected Value Loss in Treatment 4

	Mean	Std. Error
Male	4.91	0.72
	Female	
Non-Low	6.52	1.30
Low	8.50	2.57

TABLE 26. Session Attendance After Confirmation

Gender		
	Proportion	N
Male	0.79	217
Female	0.68	243
Total	0.73	460
Females by Predicted Phase		
	Proportion	N
Not Low	0.72	141
Low	0.62	102
Total	0.68	246

CHAPTER IV

GENDER DIFFERENCES IN A MARKET WITH RELATIVE PERFORMANCE FEEDBACK.

4.1 Introduction

Firms and academic institutions or grant giving agencies use tournament contests to promote individuals or allocate funds to different applicants, employees, and students. In many of these contests, individuals cannot view the quality of other competitors. The decision to enter a certain type of tournament (or firm) is largely based on subjective estimates of the probability of success. The only individuals who have good information as to the distribution of the quality of tournament contestants are the judges or evaluators that choose the winners of these contests.

In laboratory settings, where information regarding the quality of competitors is not available to potential contest entrants, it has been found that females are less likely to enter such tournament contests (Niederle & Vesterlund 2007). This gender difference in the selection into tournament contests has been used to explain the low number of females in top level management positions. A recent experiment has shown that females and males select into tournament contests in a similar manner when they are fully informed of their relative rank and the distribution of performance for all potential competitors (Chapter III). This result suggests that males and females may choose to compete in a similar fashion when participating in a labor market where relative performance information is available. Professional athletes competing in individual sports are a great test case for the effects of this information since they have access to information about the quality of their potential competitors. These athletes structure their competition schedules with fairly accurate information about the quality of competition they may face in a given contest. Based on this information, athletes may choose to enter more difficult or easier tournament contests during the competitive season. This type of scheduling and choice behavior is used to schedule tennis tournament entrants by the International Tennis Federation (ITF).

The ITF is a governing body designed to create competition opportunities and keep track of entry and results for tennis players who are trying to play for a living. For females, ITF tournaments usually involve players who are ranked from about 100 to 1,250 (or unranked) in the world. These tournaments are used as a stepping stone to qualify for the Women's Tennis Association (WTA) tournaments, which involve players ranked in the top 100 in the world and pay substantially more. For males, the ranking of individuals competing in ITF tournaments is much more diverse as players tend to be ranked from 200 to 2000 in the world. Men use these ITF tournaments to improve their world ranking and eventually qualify to compete in the more lucrative tournaments organized by the Association of Tennis Professionals (ATP).

These institutions are structured to allow individuals to enter and compete in tournaments with as much knowledge as possible about potential competitors because the institutions are designed to provide tournaments with good matches. At the ITF level, individuals provide preference ordering of the tournaments they would like to enter on a given week. These preference orderings are publicly available and, by the time a deadline approaches, individuals are best served by choosing the ordering that maximizes the expected value of their efforts. At the ATP and WTA level, the preference ordering is less of a factor as there are fewer tournaments that top level individuals enter in any given week. Though there are fewer tournaments in any given week, the prize money awarded at these tournaments is significantly higher, as are the world ranking points.[1]

By examining professional tennis players, I assess whether males and females choose to compete in a similar fashion in these competitive arenas. One would expect that the females attempting to compete in a sport for a living would be more competitive than those who do not. Given that the labor market for professional tennis players provides plenty of relative performance data, a plausible hypothesis would be that females and males choose to compete similarly in the sporting world of professional tennis. These are the most competitive tennis players for each gender; thus, one would expect similarities in how they choose to compete and how they respond to performance feedback. By using lagged performance, one can observe if females and males respond to performance feedback in a similar manner.

These tennis tournaments also provide a unique opportunity to examine potential differences in relative performance feedback using real world data as opposed to data gathered in an experimental setting. Response to performance feedback can have mixed effects. Casas-Arce and Martínez-Jerez (2009) note that it is possible that performance feedback can lower efforts in multiple period settings when agents have heterogeneous abilities. In such settings, higher ability individuals exert less

[1]World ranking points are accumulated over a 12 month rolling calendar basis. The higher the point total the better the rank of an individual.

effort than without performance feedback and lower ability individuals may find it optimal to choose to exert less effort knowing that they cannot perform better than the high ability individuals (these choices may very well be optimal). In an economic experiment concerning feedback and its effect on effort, it was found that feedback does reduce effort for different forms of compensation such as piece rates and tournaments (T. Eriksson, A. Poulsen & M.C. Villeval 2009). But in other settings performance feedback has led to more efficient choices. This is evident in studies exploring the effects of feedback for students. Azmat and Iriberri (2009) find that grades improved for the entire distribution of students when they were given relative performance feedback. Bandiera et al. (2008) find that feedback not only improved test scores, but there were also no negative performance effects from feedback.

In Chapter III, I showed that relative performance feedback leads to more efficient choices by both males and females. Previous work concluded that a preference difference for competition led to gender differences in choices for competitive environments (Niederle & Vesterlund 2007), but this result shows that this choice to compete is not driven strictly by confidence nor preferences for or against competition. Consequently relative performance feedback may be valuable in moving more high ability females towards more competitive settings. But an important question becomes: does the provision of relative performance feedback lead females towards more competitive settings? Or do males and females respond differently to relative performance feedback when choosing to enter tournaments that may affect their livelihoods? To answer these questions, I examine how performance in previous tournaments affects choices by females and males to enter competitive tournaments.

4.2 Background for Tennis Data

The major reason for using ITF tournaments is due to the structure the organization uses in scheduling tournament attendance. Each week there are multiple tennis tournaments that an athlete may attend. In 2008, there were over four-hundred ITF tennis tournaments for females with total prize money of ten-million dollars as well as over one-hundred other tournaments including Federation Cup tennis and WTA tournaments. The prize money for ITF tournaments for women is capped at one-hundred thousand dollars per tournament for females, and all WTA tournaments have greater prize money amounts.

TABLE 27. WTA Points Table

Round	W	F	SF	QF	R16	R32	R64	R128
Grand Slams	2000	1400	900	500	280	16	100	5
Premier 96S	1000	700	450	250	140	80	50	5
Premier 64S	1000	700	450	250	140	80	5	
Premier 32S	470	320	200	120	60	1		
Int. 32S	280	200	130	70	30	1		
ITF 100K+H	150	110	80	40	20	1		
ITF 100K	140	100	70	36	18	1		
ITF 75K+H	130	90	58	32	16	1		
ITF 75K	110	78	50	30	14	1		
ITF 50K+H	90	64	40	24	12	1		
ITF 50K	70	50	32	18	10	1		
ITF 25K	50	34	24	14	8	1		
ITF 10K	12	8	6	4	1			
	Points allocated as you lose round.							
2009 Prize Money example. Indian Wells.								
Prize money	605K	295K	148K	75K	40K	21K	11.5K	7K

Surprisingly, at the ITF level, males receive significantly less prize money than females. In 2008 there were over five-hundred ITF tennis tournaments for males with total prize money of six-million dollars as well as over one-hundred other tournaments including Davis Cup tennis and ATP tournaments which pay significantly more than the ITF tournaments. The prize money for ITF tournaments for men is capped at fifteen-thousand dollars per tournament. Thus, the ITF runs lower level tournaments for the males and the prize money differences between the types of tournaments that these males could choose from range between ten-thousand to fifteen-thousand dollars. Consequently there are slight institutional differences between males and females which seem to lead to greater pecuniary outcomes for females competing at lower levels. Examining the difference in world ranking points for females, in Table 27, compared to males, in Table 28, it becomes evident that the prize spread for winning is greater for males in terms of world ranking points. Such differences in point spreads between winning and losing may help explain Paserman's (2007) result that women see a drop in performance in decisive sets when compared to men.

For each week that an athlete plans to compete, he or she must provide a preference ranking for the tournaments he or she is interested in attending. For example, if there are four tournaments in a given week and an individual is interested in potentially attending one or two of them, then they must submit their preferences between the two tournaments by ranking them. Three weeks prior to the start of the tournament the preference ordering becomes closed. Individuals of the highest world rank get their first preference so long as there is space in the tournament. If there is no space in the preferred tournament then the individual gets into the

TABLE 28. ATP Points Table

Round	W	F	SF	QF	R16	R32	R64	R128
Grand Slams	2000	1250	720	360	180	90	45	10
Tour Finals (1.5K)	+500	+400	+200					
Tour Masters	1000	600	360	180	90	45	10(25)	(10)
Tour 500	500	300	180	90	45	(20)		
Tour 250	250	150	90	45	25	(10)		
Challenger 125K+H	125	75	45	25	10			
Challenger 125K	110	65	40	20	9			
Challenger 100K	100	60	35	18	8			
Challenger 75K	90	55	33	17	8			
Challenger 50K	80	48	29	15	7			
Challenger 35K+H	80	48	29	15	6			
Futures 15K+H	35	20	10	4	1			
Futures 15K	27	15	8	3	1			
Futures 10K	18	10	6	2	1			
Points allocated as you lose round.								
2009 Prize Money example. Indian Wells.								
Prize money	605K	295K	148K	75K	40K	21K	11.5K	7K

next preferred tournament conditional on there being space in that tournament. Individuals' tournament preference orderings are based on a number of characteristics relating mainly to prize money, world ranking points, and probability of winning matches. An individual may choose not to compete in a tournament after already entering. To do this, they must pay a nominal fine of about $150 dollars for ITF tournaments and up to a few thousand for higher level tournaments. At higher level tournaments, exemptions are made for minor injuries. Consequently, it is common to see athletes claim injury as opposed to paying the fine.[2]

[2]This would likely change if the size of the fine would decrease or the cost of a physician would increase.

In the data set being used for the female portion of this study, I examine the choices made by about 3,500 females for about 600 tournaments during an approximate 55 week period. For the males, I examine 3,500 males for about 1000 tournaments for about 45 weeks. The preference ordering is of interest, but very few men choose to order beyond their second choice. In comparison, females seem to choose a lot more second or third or fourth tournaments. These gender differences are shown in Table 29.

TABLE 29. Tournament Preference Observations

Preference	Females.	Males
1	64,445	45,555
2	24,821	11,007
3	14,014	5,785
4	8,604	3,636
5	5,494	2,458
6	3,332	1,710
Total	120,713	70,151

Previous economic literature has used the outcomes from professional tennis tournaments to examine the incentive effects of prizes in tournaments (T. Lallemand, R. Plasman & F. Rycx 2008, Keith F. Gilsdorf & Vasant Sukhatme 2008*a*). As well, professional tennis player data have been used to examine sequential elimination tournament effects (K.F. Gilsdorf & V.A. Sukhatme 2008*b*). These other studies examined the most elite tennis players possible. This study differs because I examine a much broader group of tennis players and because I am able to link performance data with competitive environment selection. This preference ordering of tournaments is not entirely possible at the higher level tournaments because there are far fewer tournaments occurring at the same time. Thus, at times individuals will have a top priority to enter a large tournament such as the Australian Open, a time when there are almost no other possible choices.

Despite these issues, these data provide a unique opportunity to observe both men and women entering competitive environments in a career-type setting. Though these competitors may not represent the average business professional, the opportunity to study such an environment may lead to greater insights into how individuals may make choices in more common environments.

4.3 Tournament Selection

One interesting part part of the tennis labor market is that all competitors receive a noisy signal of the quality of potential opponents in the form of a world ranking. The quality of a tennis player, $Q(x_i)$, is a function of a number of inputs (x_i), such as individual talent, coach quality, hours practiced, allocation of hours towards specific skills, mental toughness, emotional control... specific to individual i. The signal used by most to determine quality is a sum of an individual's world ranking points (R_i), but this signal includes an exogenous error component as shown below:

$$R_i = Q(x_i) + \varepsilon_i$$

Suppose that two individuals, i and j, decide to compete to see who is the better tennis player. The probability that i beats j is then as follows:

$$\begin{aligned} \text{Prob}(Q(x_i) > Q(x_j)) &= \text{Prob}(R_i - R_j > \varepsilon_j - \varepsilon_i) \\ &= \text{Prob}(\varepsilon_i - \varepsilon_j > R_j - R_i) \end{aligned}$$

It is because of this noisy signal of quality that individuals, i and j, find it necessary even to play the game. If quality was visible and measured without error, or without luck, then competition between these individuals would be unnecessary. This uncertainty is why sports have audiences and teams and athletes have supporters. Even when there is a winner or a loser, one cannot distinguish whether the victory came from the difference in quality between opponents or from an exogenous error component often described as luck. A major difference between sports and other labor markets is that there is usually a clearer signal of quality in sports markets due to numerous different forms of relative ranking methodologies. Such rankings are non-existent in the majority of other labor markets. It is likely that the cost of creating a measure to find the best copy writer in the world does not outweigh the benefit, though it very much matters when it comes to hitting a yellow fuzzy ball.

Consider an individual who is examining a set of N possible tournament contests to enter. Each tournament n, is such that $n \in \{1, 2, 3..., N\}$. In tennis the individual receives a signal regarding the quality of potential opponents by examining the preferences and rankings of all individuals that are considering entering the tournament. The individual may use a moment (e.g. mean) of the distribution of world rank points for competitors entering the tournament. This moment allows each competitor to gauge the probability of winning a match against the average tournament entrant. Let $\bar{R}_n$ represent the mean ranking points of potential opponents entering tournament n, such that $\bar{R}_n = \bar{Q}_n + \varepsilon_n$. Assuming independence between rounds, this implies that the probabilities of winning rounds in a sequential match play tournament are as follows:

Probability of winning first round	$\text{Prob}(\varepsilon_i - \bar{\varepsilon}_n > \bar{R}_n - R_i)$
Probability of winning second round	$\text{Prob}(\varepsilon_i - \bar{\varepsilon}_n > \bar{R}_n - R_i)^2$
Probability of winning k round	$\text{Prob}(\varepsilon_i - \bar{\varepsilon}_n > \bar{R}_n - R_i)^k$

These probabilities allow an individual to calculate the expected value of entering each of the available tournaments. For simplicity, I am ignoring the non-random scheduling of opponents for each round and I am ignoring other compensating differentials that may lead an individual to choose a tournament that does not maximize expected value from cash prizes and ranking points. In the empirical model, such potentially confounding factors will be included. I take as given, the effects upon optimal effort from prize differences, since this subject has been previously explored using tennis data (Gilsdorf & Sukhatme 2008*a*). I assume that all individuals choose optimal effort levels and consider the prize spread between winning and losing in choosing those effort levels. This prize spread and prize level is partly a reflection of the total tournament prize money; thus, the prize money will be used as a proxy for the monetary benefit from competition.

Using these win probabilities for each round and prize money and world rank point values for winning each round an individual can compare the expected values from competing in each of the different tournaments. Individuals should then rank order tournaments according to these subjectively calculated expected values.[3] It

[3]For simplicity I am ignoring the strategic implications of the publishing of the preference ordering for tournaments by individuals. At the time of the deadline, I assume that all individuals choose

is this type of selection that will be considered in this study. The following section provides a brief analysis of the data collected for both men and women tennis players.

4.4 Tournaments and Performance

Data collection for tournament preference orderings for professional female tennis players began in March 2009. Females rank order up to six tournaments each week that they consider entering. There are many players who only enter one tournament, but these individuals are less likely to be travelling professionals; thus, a large portion of the sample consists of a single preference of 1. Altogether, there are 120,713 observations of tournament preferences by females. The tabulations for all the preference ordering data are provided in Table 29.

TABLE 30. Summary statistics for women.

Variable	Mean	Std. Dev.	N
WeekOfTourney	32.58	16.84	21790
Ranking	827.34	408.68	123244
Age	22.32	3.47	115157
Tourn. Total Games Won	25.78	16.32	18989
Tourn. Total Games Lost	14.48	10.85	18989
Prize Money	260116.81	1777635.23	117458
TourneyRank Avg.	323.31	101.13	97048
Priority	1.97	1.34	120713
HomeField	0.06	0.24	128203
Tourn. Match Wins	0.99	1.29	18989
LagTourn. Total Games Won	3.72	10.98	128203

to reveal themselves such that they maximize expected value. The assumption is that any one individual will not change the average rank of tournament competitors.

To go along with this tournament entry data, I have also collected performance data for all individuals who entered any professional tournaments during the same period. The summary statistics for the performance of females and tournament preferences are shown in Table 30. The average total purse per preference observation is $260,000. The average world ranking of competitors is 827. The average rank per tournament is 323.

For this study, data collection for tournament preference ordering for professional male tennis players began in July 2009. Though males also rank order up to six tournaments each week, there are many individuals who choose to provide a preference for only one tournament. However, about 38% of top priority choices for females also include an alternative. For males this percentage is much lower, at 24%. This difference suggests that males must see less need to provide possible secondary choices to compete. Altogether during this time there are approximately 70,000 observations of tournament preference choices by males.

TABLE 31. Summary statistics for men.

Variable	Mean	Std. Dev.	N
WeekOfTourney	35.68	17.4	27890
Ranking	586.27	556.9	21003
Age	23.89	4.13	27639
Tourn. Total Games Won	26.45	16.93	18860
Tourn. Total Games Lost	16.31	11.88	18860
Prize Money	402671.48	2493353.52	84314
Avg. Rank Tourney	836.98	337.41	80352
Priority	1.74	1.26	70151
HomeField	0.78	0.41	88331
Tourn. Match Wins	1.01	1.31	18860
LagTourn. Total Games Won	5.26	13.28	88638

The summary statistics for the performance of males and tournament preferences are shown in Table 31. The average total purse per preference observation is \$403,000. The average world ranking of competitors is 586. The average rank per tournament is 836. As Figure 12 shows, the prize money tournaments for males and females are scheduled at similar times. Males do seem to have a larger number of tournaments between \$100,000 to \$4.5 million when compared to females, but on the lower spectrum females have more tournaments between \$15000 to \$100,000.

For the lower prize money tournaments, the data had to be collected manually. This implies that the lower ranked individuals do not have their priority choice observed for the entire sample. The female sample consists of more weeks than the male sample. This difference in the samples is best shown by the ranking of individuals entering tournaments as individuals with a ranking greater than 500 had to be recorded manually. This gender specific sample difference is best observed in Figure 13.

The number of matches that a competitor wins at a tournament is a measure of success. I condense the data to tournament-specific values as opposed to examining results round by round as this allows for a clearer comparison of tournament specific factors that may determine success. In Table 32, I provide the log-linear regression estimates using both the random and fixed effects for individuals to observe which factors are positively related to success. The dependent variable to measure success is the log of match wins. I find that estimates of the systematic effects of some factors, such as hospitality (a categorical variable determining whether expenses were covered by the tournament), individual ranking, and previous performance from prior weeks are very sensitive to the assumed error structure.

For females, home field advantage has a a large and significant impact on individuals being able to win. *Homefield* is a categorical variable indicating whether the tournament is hosted in an individual's home country. *Top priority* indicates that a tournament was an individual's top choice and surprisingly it has a negative effect on performance, which may indicate that females prefer to enter more difficult tournaments. Not surprisingly, the average rank of players entered in a tournament is positively correlated with success as the higher the rank of potential competitors, the lower the quality of the potential competitors, the more likely an individual can be successful.

The potential factors affecting male success in tournaments are shown in Table 32. The estimated effect of previous performance is very sensitive to error assumptions. As with females, the average rank of potential competitors is positively correlated with performance. Unlike females, males' top choices for tournaments are positively correlated with tournament success. This result suggests that males may be better at choosing which tournaments to enter though one should note that females typically have a larger choice set to choose from.

These performance regressions suggest that some factors such as the rank of potential competitors and competing in one's home country may help both a female and male competitor to be successful. Thus, these factors should also entice individuals to enter such tournaments. On the otherhand, males and females see performance differences when they enter in tournaments that were their top choice,

TABLE 32. Log of Match Wins in Tourney

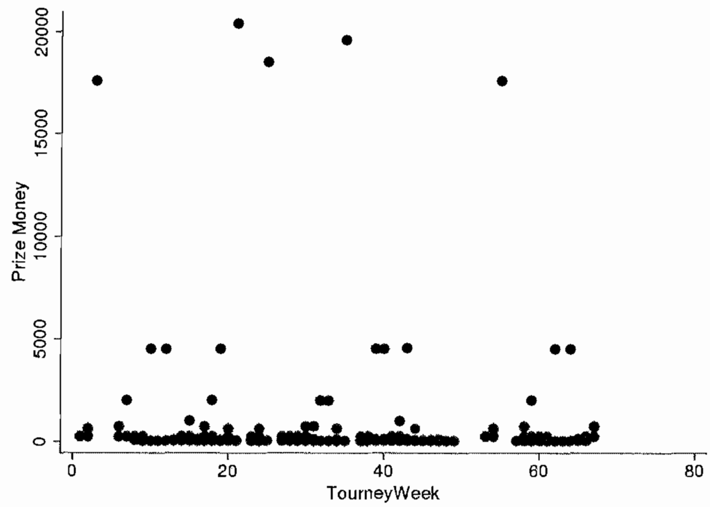
VARIABLES	(1) Female RE	(2) Female FE	(3) Male RE	(4) Male FE
Log_PrizeMoney	-0.110	-0.130	-0.128	0.168
	(0.041) ***	(0.041) ***	(0.185)	(0.224)
hospitality	-0.016	-0.508	-0.022	0.190
	(0.159)	(0.170) ***	(0.153)	(0.175)
HomeField	0.231	0.350	-0.002	-0.005
	(0.094) **	(0.109) ***	(0.073)	(0.101)
Log_TourneyRankAvg	2.898	3.525	2.668	3.781
	(0.191) ***	(0.203) ***	(0.236) ***	(0.277) ***
Log_ranking	-1.491	0.539	-2.391	1.911
	(0.066) ***	(0.161) ***	(0.077) ***	(0.267) ***
Log_Age	-0.668	16.297	0.284	93.357
	(0.263) **	(16.837)	(0.258)	(29.561) ***
TopPriority	-0.444	-0.538	0.241	0.254
	(0.099) ***	(0.114) ***	(0.118) **	(0.134) *
Log_WeekLagTournMatchWins_1	0.040	-0.005	0.092	-0.054
	(0.010) ***	(0.010)	(0.010) ***	(0.011) ***
Log_WeekLagTournMatchWins_2	0.017	-0.013	0.025	-0.075
	(0.010)	(0.010)	(0.011) **	(0.011) ***
Week	Y	Y	Y	Y
Surface controls	Y	Y	Y	Y
Observations	9,119	9,119	9,821	9,821
Number of id_p	704	704	1,944	1,944
r2_o	0.0775	0.00339	0.145	0.0138
r2_w	0.0626	0.0957	0.00125	0.0503
r2_b	0.159	0.00175	0.389	0.0431

Standard errors in parentheses

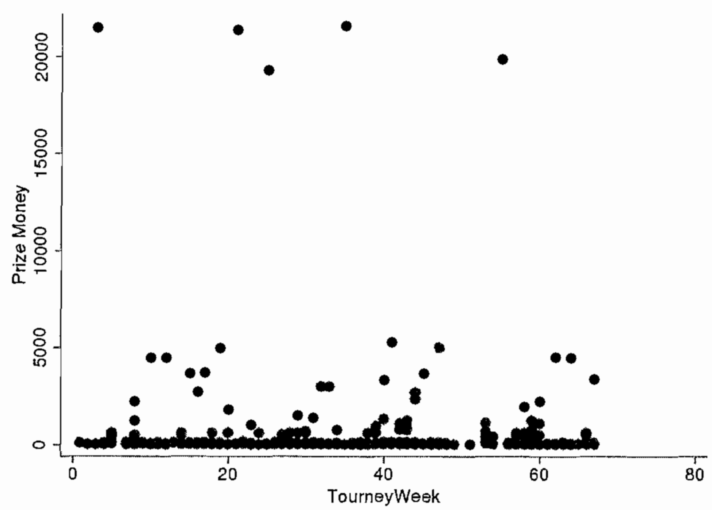
*** p<0.01, ** p<0.05, * p<0.1

suggesting that males and females choose to enter tournaments differently. To explore this issue, in the next session, I provide hazard function coefficient estimates for both males and females to see if and how they may be choosing to enter into tournaments differently.

FIGURE 12. Tournaments and Prize Money

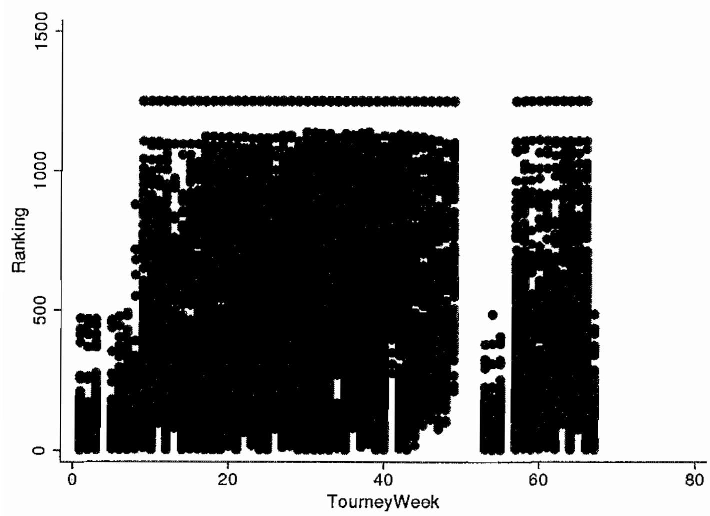


Female Tournaments

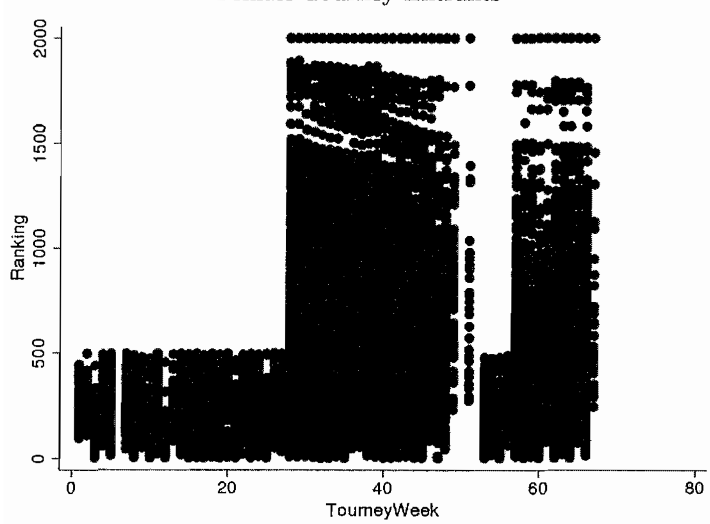


Male Tournaments

FIGURE 13. Rankings and Top Priority



Female Tourney Entrants



Male Tourney Entrants

4.5 Gender Differences for Tournament Entry

To examine if females and males choose to enter tournaments differently I estimate a hazard model where failure is dictated by tournament entry into any tournament. A hazard model should be appropriate for this type of comparison as I am interested in comparing how men and women choose to compete. Thus a hazard model with failure being associated with tournament entry provides an estimation of which factors increase or decrease the chance of a failure occurring, which is an occurrence of competition. An examination of these estimates should allow me to observe whether females and males choose to compete differently.

I estimate the model separately for males and females. I find that males and females seem to use very different criteria to enter. The results of the estimation for females are shown in Table 33. Females are more likely to enter into a tournament if it is in their home country (*HomeField*) and if they have a lower ranking. The prize money effects are for entry are negligible or are not apparent for tournaments.

The two measures of *WeekLagTournMatchWins_1* and *WeekLagTournMatch-Wins_2* measure how well individuals performed in a tournament in the previous week (1) or two (2). Most importantly, being more successful in the previous two tournaments leads to a lower chance of tournament entry. So positive performance feedback from a competition leads to less subsequent competition by females.

The male hazard model coefficients are shown in Table 34. Three major differences become evident when comparing males' and females' entries into tournaments in this setting:

TABLE 33. Female Tourney Entry

VARIABLES	(1) _t	(2) _t
PrizeMoney	0.000	0.000
	(0.000)	(0.000)
hospitality	-0.094	-0.094
	(0.044) **	(0.046) **
HomeField	0.204	0.204
	(0.029) ***	(0.029) ***
TourneyRankAvg	-0.000	-0.000
	(0.000)	(0.000)
ranking	-0.004	-0.004
	(0.000) ***	(0.000) ***
Age	-0.009	-0.009
	(0.003) ***	(0.003) ***
WeekLagTournMatchWins_1	-0.032	-0.032
	(0.008) ***	(0.009) ***
WeekLagTournMatchWins_2	-0.021	-0.021
	(0.008) **	(0.008) **
Surface controls	Y	Y
Observations	11,286	11,286
risk	33659	33659
N_fail	8700	8700
N_sub	710	710
chi2	3640	2082
N_clust		710

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

1. Males are more likely to enter a tournament when housing is covered by tournament organizers shown by the *hospitality* variable. Females do not respond to such incentives.

2. Men respond to how competitive the field is at the tournament. A higher average ranking of tournament entrants (*TourneyRankAvg*) makes for easier competition and also makes it more likely that males will enter the tournament. This average ranking has no effect on female tournament entry though it does positively affect performance for females.

3. Most interestingly, the effect of relative performance feedback is completely different for males than for females. If males do well in a tournament in the previous week then they are more likely to enter into a tournament. On the other hand females are less likely to enter a tournament if they perform well in tournaments that occur two weeks prior to the tournament they are entering.

TABLE 34. Male Tourney Entry

VARIABLES	(1) _t	(2) _t
PrizeMoneyNoCap	0.051	0.051
	(0.004)	(0.005)
	***	***
hospitality	0.436	0.436
	(0.044)	(0.054)
	***	***
HomeField	0.214	0.214
	(0.022)	(0.029)
	***	***
TourneyRankAvg	0.003	0.003
	(0.000)	(0.000)
	***	***
ranking	-0.001	-0.001
	(0.000)	(0.000)
	***	***
Age	-0.037	-0.037
	(0.003)	(0.004)
	***	***
WeekLagTournMatchWins_1	0.113	0.113
	(0.010)	(0.014)
	***	***
WeekLagTournMatchWins_2	0.015	0.015
	(0.011)	(0.015)
Surface controls	Y	Y
Observations	9,463	9,463
risk	64599	64599
N_fail	8535	8535
N_sub	1935	1935
chi2	3192	2130
N_clust		1935

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

4.6 Concluding Remarks

I find that males and females behave differently when it comes to selecting into competitive environments even in a market with good relative performance information. Previous findings (Chapter III) would suggest that relative performance feedback may be enough to make males and females enter into competitive situations in a similar fashion. In examining tournament entry by professional tennis players I have found that good relative performance feedback may not be enough.

Females are less likely to remain competitive after being successful in recent competitions, while males get more competitive with success in recent competitive tournaments. Males seem to base part of the decision about whether to compete on the strength of whom they are competing against. In contrast, females do not make decisions on whether to compete based on how difficult the competition may be.

It seems that feedback for competition may not necessarily be a good thing to entice females to compete. Even in a high end of the distribution feedback has different effects for females and males. It may be the case that females and males should have to pay for feedback. It would seem that males would be willing to pay for such information while females would not. It may be the case that females do not necessarily have distaste for competition, but do have a distaste for publicly being shown to be better than one another.

CHAPTER V

CONCLUSION

The preceding chapters have shown that information effects matter a great deal for individual decision making about competition. Biology also plays a role and biological functions such as hormonal fluctuations may lead to large changes in the gender difference for for competition. Though in a laboratory setting it seemed that good information could remove the effects of biological differences for entry into competitive tournaments, it seems that this is not entirely the case. Even in markets with limited ambiguity, consisting of very competitive individuals, gender differences for entry into competitive situations still remain. Furthermore the response to positive and negative performance feedback leads to very different behaviors between the genders.

Though a lack of ambiguity can change some behaviors of agents at the margin, it seems that good information alone is not enough to change gender differences for competition. It may be that these differences are socially optimal as in repeated similar situations, under similar contexts, men and women compete differently. Using the measure of ambiguity aversion, as discussed in Chapter II, one can test as to the role that ambiguity plays in individual decisions to compete. Such investigations will provide greater insight into the mechanisms driving gender differences in labor markets. This could lead to different policies in labor markets as firms may be better served by providing more information about the quality of applicants applying for positions. Such policies may remove ambiguity concerning the competitive applicant pool, which may lead more high ability women to compete for such positions. The goal is to find the mechanisms leading to higher ability individuals entering competitive situations. Given the gender differences explored here, it seems that we have moved closer to understanding the gender differences for competition.

BIBLIOGRAPHY

Aden, U., C. Jung-Hoffmann, and H. Kuhl. 1998. "A randomized cross-over study on various hormonal parameters of two triphasic oral contraceptives." *Contraception*, 58(2): 75–81.

Anderson, FD, and H. Hait. 2003. "A multicenter, randomized study of an extended cycle oral contraceptive." *Contraception*, 68(2): 89–96.

Azmat, G., and N. Iriberri. 2009. "The Importance of Relative Performance Feedback Information: Evidence from a Natural Experimento using High School Students." *Working Paper.*

Bandiera, O., V. Larcinese, and I. Rasul. 2008. "Blissful Ignorance? Evidence from a Natural Experiment on The Effect of Individual Feedback on Performance." Mimeo.

Barber, B.M., and T. Odean. 2001. "Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment." *The Quarterly Journal of Economics*, 116(1): 261–292.

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

Becker, G.M., M.H. DeGroot, and J. Marschak. 1963. "Stochastic Models of Choice Behavior." *Behavioral Science*, 8(1): 41–55.

Becker, S.W., and F.O. Brownson. 1964. "What price ambiguity? Or the role of ambiguity in decision-making." *The Journal of Political Economy*, 62–73.

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

Bethea, C.L., N.Z. Lu, C. Gundlah, and J.M. Streicher. 2002. "Diverse Actions of Ovarian Steroids in the Serotonin Neural System." *Frontiers in Neuroendocrinology*, 23(1): 41–100.

Borghans, L., B.H.H. Golsteyn, J.J. Heckman, and H. Meijers. 2009. "Gender Differences in Risk Aversion and Ambiguity Aversion." *Journal of the European Economic Association*, 7(2-3): 649–658.

Brown, C. 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.

Casas-Arce, Pablo, and F. Asís Martínez-Jerez. 2009. "Relative Performance Compensation, Contests, and Dynamic Incentives." *Manage. Sci.*, 55(8): 1306–1320.

Charness, G., and Uri Gneezy. 2007. "Strong Evidence for Gender Differences in Investment." *SSRN*.

Chen, Yan, P. Katuscak, and E. Ozdenoren. 2005. "Why Can't a Woman Bid More Like a Man?" *Working Paper, Ann Arbor*, 1001: 48109–1220.

Chipman, J.S. 1960. "Stochastic choice and subjective probability." *Decisions, Values, and Groups: Reports*, 70.

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, E.M., and S.J. Segal. 1999. *Is menstruation obsolete?* Oxford University Press, USA.

Dohmen, T.J., and A. Falk. 2007. "Performance pay and multi-dimensional sorting: productivity, preferences and gender." *Institute for the Study of Labor (IZA); CESifo (Center for Economic Studies and Ifo Institute for Economic Research)*.

Ellsberg, D. 1961. "Risk, ambiguity, and the Savage axioms." *The Quarterly Journal of Economics*, 643–669.

Eriksson, T., A. Poulsen, and M.C. Villeval. 2009. "Feedback and incentives: experimental evidence." *Labour Economics*.

Fink, G., BE Sumner, R. Rosie, O. Grace, and JP Quinn. 1996. "Estrogen control of central neurotransmission: effect on mood, mental state, and memory." *Cell Mol Neurobiol*, 16(3): 325–44.

Fox, C.R., and A. Tversky. 1995. "Ambiguity Aversion and Comparative Ignorance." *Quarterly Journal of Economics*, 110(3): 585–603.

Gigliotti, G., and B. Sopher. 1996. "The testing principle: Inductive reasoning and the Ellsberg paradox." *Thinking & Reasoning*, 2(1): 33–49.

Gilsdorf, Keith F., and Vasant Sukhatme. 2008*a*. "Tournament incentives and match outcomes in women's professional tennis." *Applied Economics*, 40(18): 2405–2412.

Gilsdorf, K.F., and V.A. Sukhatme. 2008*b*. "Testing Rosen's Sequential Elimination Tournament Model: Incentives and Player Performance in Professional Tennis." *Journal of Sports Economics*, 9(3): 287.

Gneezy, U., and A. Rustichini. 2004. "Gender and Competition at a Young Age." *American Economic Review*, 94(2): 377–381.

Gneezy, U., K.L. Leonard, and J.A. List. 2009. "Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society." *Econometrica*, 77(5): 1637–1664.

Gneezy, U., M. Niederle, and A. Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*, 118(3): 1049–1074.

Grossman, P.J., and C. Eckel. 2003. "Men, Women and Risk Aversion: Experimental Evidence." *Unpublished manuscript.*

Gunther, C., N.A. Ekinci, C. Schwieren, and M. Strobel. 2008. "Women can't jump?–An experiment on competitive attitudes and stereotype threat." Mimeo.

Gupta, N.D., A. Poulsen, and M. Villeval. 2005. "Do (wo) men prefer (non-) competitive jobs." GATE Working Paper.

Hakes, JK, and WK Viscusi. 2004. "Dead Reckoning: Demographic Determinants of the Accuracy of Mortality Risk Perceptions." *Risk Analysis*, 24(3): 651–664.

Halevy, Y. 2007. "Ellsberg Revisited: An Experimental Study." *Econometrica*, 75(2): 503–536.

Harbaugh, W.T., K. Krause, and L. Vesturlund. 2002. "Risk Attitudes of Children and Adults: Choices Over Small and Large Probability Gains and Losses." *Experimental Economics*, 5: 53–84.

Heath, C., and A. Tversky. 1991. "Preference and belief: Ambiguity and competence in choice under uncertainty." *Journal of Risk and Uncertainty*, 4(1): 5–28.

Holt, C.A., and S.K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review*, 92(5): 1644–1655.

Ichino, A., and E. Moretti. 2009. "Biological Gender Differences, Absenteeism and the Earnings Gap." *American Economic Journal: Applied Economics*, 1(1): 183–218.

Jirjahn, U., and G. Stephan. 2004. "Gender, piece rates and wages: evidence from matched employer-employee data." *Cambridge Journal of Economics*, 28(5): 683–704.

Kagel, J.H., and A.E. Roth. 1995. *The Handbook of Experimental Economics.* Princeton University Press Princeton, NJ.

Knight, F.H. 1921. *Risk, Uncertainty and Profit.* Houghton Mifflin Company.

Kosfeld, M., M. Heinrichs, P.J. Zak, U. Fischbacher, and E. Fehr. 2005. "Oxytocin increases trust in humans." *Nature*, 435(7042): 673–676.

Lallemand, T., R. Plasman, and F. Rycx. 2008. "Women and Competition in Elimination Tournaments: Evidence From Professional Tennis Data." *Journal of Sports Economics*, 9(1): 3.

Lazear, E.P., and S. Rosen. 1981. "Rank-order tournaments as optimum labor contracts." *The Journal of Political Economy*, 89(5): 841.

Lichtenstein, S., P. Slovic, B. Fischhoff, M. Layman, and B. Combs. 1978. "Judged frequency of lethal events." *Journal of Experimental Psychology: Human Learning and Memory*, 4(6): 551–578.

Malison, R.T., L.H. Price, R. Berman, C.H. van Dyck, G.H. Pelton, L. Carpenter, G. Sanacora, M.J. Owens, C.B. Nemeroff, N. Rajeevan, et al. 1998. "Reduced brain serotonin transporter availability in major depression as measured by [123I]-2β-carbomethoxy-3β-(4-iodophenyl) tropane and single photon emission computed tomography." *Biological Psychiatry*, 44(11): 1090–1098.

Matsushita, R., D. Baldo, B. Martin, and S. Da Silva. 2007. "The biological basis of expected utility anomalies."

McEwen, B. 2002. "Estrogen Actions Throughout the Brain." *Recent Progress in Hormone Research*, 57(1): 357–384.

Moore, E., and C. Eckel. 2003. "Measuring Ambiguity Aversion. Virginia Tech."

Niederle, M., C. Segal, and L. 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.

Paserman, M. Daniele. 2007. "Gender Differences in Performance in Competitive Environments: Evidence from Professional Tennis Players." Institute for the Study of Labor (IZA) IZA Discussion Papers 2834.

Pearson, M., and B.C. Schipper. 2009. "Menstrual cycle and competitive bidding." *MPRA Paper.*

Platt, ML, and SA Huettel. 2008. "Risky business: the neuroeconomics of decision making under uncertainty." *Nat Neurosci*, 11(4): 398–403.

Polachek, S.W. 1981. "Occupational self-selection: A human capital approach to sex differences in occupational structure." *The Review of Economics and Statistics*, 60–69.

Saha, A. 1993. "Expo-Power Utility: A Flexible Form for Absolute and Relative Risk Aversion." *American Journal of Agricultural Economics*, 75: 905–905.

Savage, L.J. 1972. *The Foundations of Statistics.* Dover publications.

Schmeidler, D. 1989. "Subjective probability and expected utility without additivity." *Econometrica: Journal of the Econometric Society*, 571–587.

Speroff, L., and M.A. Fritz. 2005. *Clinical Gynecologic Endocrinology and Infertility.* Lippincott Williams & Wilkins.

Stanton, S.J., and O.C. Schultheiss. 2007. "Basal and dynamic relationships between implicit power motivation and estradiol in women." *Hormones and Behavior*, 52(5): 571–580.

Stricker, R., R. Eberhart, M.C. Chevailler, F.A. Quinn, P. Bischof, and R. 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.

Vliet, E.L. 2001. *Screaming to be heard: hormone connections women suspect and doctors still ignore.* New York, NY:M. Evans and Co.

Zak, P.J., R. Kurzban, and W.T. Matzner. 2005. "Oxytocin is associated with human trustworthiness." *Hormones and Behavior*, 48(5): 522–527.

Zethraeus, N., L. Kocoska-Maras, T. Ellingsen, B. von Schoultz, A.L. Hirschberg, and M. 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.