Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

On the Size of the Gender Difference in Competitiveness

Silvia Saccardo, Aniela Pietrasz, Uri Gneezy

MANAGEMENT SCIENCE
Vol. 64, No. 4, April 2018, pp. 1541–1554
ISSN 0025-1909 (print), ISSN 1526-5501 (online)
http://pubsonline.informs.org/journal/mnsc/

On the Size of the Gender Difference in Competitiveness

Silvia Saccardo,[a] Aniela Pietrasz,[b] Uri Gneezy[c,d]

[a] Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213; [b] University of California, San Diego, La Jolla, California 92093; [c] Rady School of Management, University of California, San Diego, La Jolla, California 92093; [d] Center for Research in Experimental Economics and Political Decision Making, University of Amsterdam, 1018 WB Amsterdam, Netherlands

Contact: ssaccard@andrew.cmu.edu (SS); aniela.pietrasz@gmail.com (AP); ugneezy@ucsd.edu (UG)

Received: May 4, 2014
Revised: July 27, 2015; April 5, 2016; August 21, 2016
Accepted: August 27, 2016
Published Online in Articles in Advance: May 1, 2017

https://doi.org/10.1287/mnsc.2016.2673

Abstract. We design a new procedure for measuring competitiveness and use it to estimate the magnitude of the gender gap in competitiveness. Before working on a task, participants choose what percentage of their payoffs will be based on a piece-rate compensation scheme; the rest of their payoff is allocated to a competitive compensation scheme. This novel procedure allows us to distinguish between 101 levels of competitiveness, as opposed to the binary measure used in the literature. Whereas the binary measure allows researchers to conclude that about twice as many men as women choose to compete (typically two-thirds versus one-third), the new procedure sheds light on the intensive margin. We find that the intensity of the preference is more extreme than the binary measure could detect. For example, we find that only one-fifth of the most competitive 25% of our participants are women, and the most competitive 10% of our participants are all men. The new procedure also allows us to study the correlation between competitiveness and parameters such as overconfidence, attitudes toward risk, and ambiguity.

History: Accepted by Gérard Cachon, behavioral economics.

Funding: This research was conducted under the Institutional Review Board [Research Approval Number 110022], and benefited from the financial support of internal funds at the Rady School of Management, University of California, San Diego.

Supplemental Material: Data and the online appendix are available at https://doi.org/10.1287/mnsc.2016.2673.

Keywords: gender • competitiveness • behavioral economics

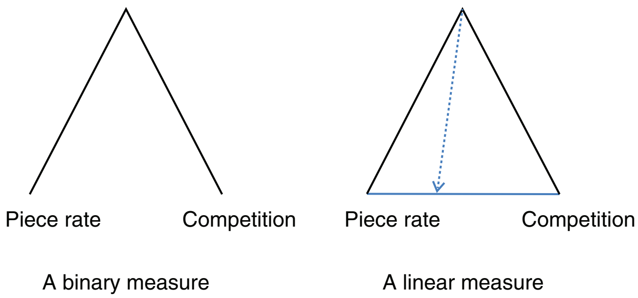
1. Introduction

In 2014, the number of female chief executive officers (CEOs) of Fortune 500 companies reached a historical high of 24 (or 4.8% of all CEOs). Although the increase in female leadership is encouraging, this number shows that the gender gap in the labor market is still large.[1] All over the world, women earn, on average, less than men in similar jobs, and in all but four countries, females account for substantially less than half of the senior positions in business and government. Strikingly, women account for their share of such positions in only 11 countries (2.4% of the world population; see Hausmann et al. 2010). Researchers have proposed a large variety of causes to explain this difference, including discrimination and differences in work–home preferences. In this paper, we focus on the magnitudes of individuals' preferences for competition.

Over the last decade, a stream of experimental research has argued that women are less competitive than men and that this difference could partially explain the differential success between men and women in the labor market (see Croson and Gneezy 2009 for a survey). This research has focused on two aspects of competitiveness. One line of literature measured participants' reactions to changes in the competitive nature of the compensation schemes and showed that, when forced into a competitive setting, women perform worse than men (e.g., Gneezy et al. 2003, Günther et al. 2010, Shurchkov 2012). A second line of literature has investigated individuals' self-selection into competitive environments. Typically, these papers ask participants to choose a compensation scheme for themselves (e.g., Niederle and Vesterlund 2007). The major result of this line of literature is that women select into competitive environments less often than men.[2]

The discussion in the evolutionary literature on the origins of gender differences in behavior is classic, going back to Darwin (1871), Bateman (1948), and Trivers (1972). A large body of literature in evolutionary biology and sociobiology documents differences in competitiveness between males and females in many different species (Knight 2002; see Dekel and Scotchmer 1999 for a model of the economic implications of such evolutionary differences). The question of how the intensity of preferences with regard to competitiveness differs between men and women is still open. One possibility is that men's and women's competitiveness levels differ, for example, because of nature, as in the evolutionary models. In particular, according to the evolutionary argument, a direct preference effect occurs because men are more likely to have

Figure 1. (Color online) Measures of Competitiveness on the Extensive and Intensive Margins



Notes. The left side represents the binary measure in which participants are asked to choose between a piece-rate compensation scheme and a change to tournament-based compensation scheme. The right side represents the linear measure in which the participants can choose a combination of the two compensation schemes.

some inherent taste for competition distinct from, for example, risk and ambiguity preferences—a "competitive spirit." This explanation would imply that the distribution of competitive spirit differs by gender.

An alternative explanation is that even if men and women have the same competitive tendency, their choice of entering a competition may differ because of indirect preference. For example, even if the initial disposition of the competitiveness of women is the same as for men, nurture can make women and men more or less likely to compete (see Gneezy et al. 2009).

To investigate these explanations, one needs a more refined measure of competitiveness than the one used in the current literature. In particular, we need to know more about the intensity of this competitive spirit by gender. The papers that study selection into competitive environments typically use the choice of incentive scheme as a binary measure of competitive behavior: participants are asked to choose whether they would like to be paid according to a piece-rate incentive scheme or a tournament incentive scheme (as presented in the left part of Figure 1). Note that in this line of research, "competitiveness" is defined as the selection of the tournament scheme. A robust finding of this line of research is that men choose competitive incentives (the tournament scheme) more often than women; hence, the "average woman" is less competitive in this context than the "average man." However, a binary measure, which focuses on competitiveness on the extensive margin (the choice of whether to enter the competition), does not allow us to make claims about the intensity of competitive preferences.

Furthermore, although fewer women than men choose to compete in these experiments, on average, around one-third of women do choose the competitive incentives. Hence, when considering the role of competitiveness in explaining the extent of the gender gap in elite labor market outcomes, these results suggest that about one-third of the women are as competitive as the competitive men. The discrepancy between the gender differences in competitiveness observed in these experiments relative to the vast gap in, for example, the number of top male and female CEOs suggests that competitiveness must play only a relatively minor role. Learning more about the distribution of competitiveness would allow us to investigate possible differences in the *intensity* of competitive preferences between men and women, providing insights on the strength of their preferences. Furthermore, if some positions require high levels of competitiveness, investigating potential differences in the fraction of men and women among highly competitive people is important.

In this paper, we introduce a measure that allows us to observe 101 different levels of competitiveness: we ask participants to choose what percentage of their compensation they would prefer to be derived from the tournament scheme and what percentage derived from the piece-rate scheme (as presented in the right side of Figure 1). This measure is similar in spirit to Gneezy and Potters' (1997) measure of risk aversion, and it allows us to measure an intensive margin of competitiveness and, therefore, to better understand the intensity of preferences for competitiveness.

Through this more refined experimental procedure, we are able to collect a richer data set that allows us to perform analyses that are more informative than the examination of averages alone, shedding light on the strength of men's and women's preferences. We measure the extent to which one individual is more competitive than another—and examine how the distribution of competitiveness differs between the populations of men and women. The differences between men and women at the upper tail of the distribution of competitiveness cannot be observed through a binary choice—conditional on choosing to compete, any two levels of competitiveness within the same binary classification will appear to be the same. Our results reveal that even conditional on choosing to compete, men exhibit a stronger preference than women. When looking at the population of men and women combined, the ratio of women to men in the distribution of competitiveness decreases as the degree of competitiveness increases. Although the women-to-men ratio in the extensive margin reveals that approximately one-third of all of the people who choose the tournament are women, the intensive-margin measure reveals that this proportion is much smaller in the upper tail of the distribution of competitiveness, where the women-to-men ratio substantially decreases. Strikingly, all of the participants in the top 10% of the distribution of competitiveness are men. Our results suggest that the distribution of competitive spirit in our sample is gender dependent.

Our refined measure of competitiveness also allows us to investigate the extent to which gender differences in beliefs and other economic preferences account for the gender gap in competitiveness. That is, we can further explore whether competitiveness differs across gender irrespective of other preferences or whether it is determined indirectly by other gender differences. We focus on three factors that may affect the extensive and intensive margin of competitiveness. First, we focus on differences in confidence (e.g., Lichtenstein et al. 1982, Soll and Klayman 2004, Niederle and Vesterlund 2007, Möbius et al. 2011, Balafoutas and Sutter 2012, Niederle et al. 2012). Second, we also investigate whether gender differences in attitudes toward risk (see, e.g., Eckel and Grossman 2008, Charness and Gneezy 2012; see Croson and Gneezy 2009 for a survey) can account for a share of the gender gap in competitiveness. Third, we explore the relationship between competitiveness and attitudes toward ambiguity, a factor that has received little or no attention in previous investigations of the gender gap in tournament entry.

The remainder of this paper is organized as follows: The next section illustrates the experimental design, providing a description of the measures of competitiveness on the intensive and extensive margin and of the additional measures collected during the experiment. Section 3 discusses the results of the measure on the extensive margin, the results of the measure on the intensive margin, and the comparison between the measures. Section 4 concludes.

2. Experimental Design

In our experiment, all participants faced the same task, and competitiveness was measured either using a binary choice between the tournament and piece-rate (extensive margin) or by choosing a linear combination of the two (intensive margin). For this purpose, we employed two treatments. Participants in both treatments chose how they wanted to be compensated for completing a ball-tossing task (Gneezy et al. 2009). The ball-tossing task involves tossing a tennis ball into a small basket 10 feet away. In the task, participants are given 10 opportunities to make successful tosses—tosses that land and stay in the basket are considered successful. Each toss must be completed underhand. We explained the task in detail to the participants at the start of each experimental session. That is, while reading the task instructions out loud, the research assistant showed the tennis ball to the participants and mimicked tossing it into a basket placed 10 feet away. To prevent a given outcome from affecting participants' beliefs about the difficulty of the task, the experimenter did not actually perform the task. Participants performed the task in private so that no other participant could observe their performance.

To exclude the possibility of a gender difference in ability in our sample, we conducted a separate between-participants test in which we measured ability by asking participants to engage in the task without letting them choose their compensation scheme. This test consisted of 84 participants (42 women) who belonged to the same participant pool as in the main experiment. These participants took part in an unrelated experiment and completed the task for no incentives. The results, reported in Online Appendix A1, reveal no gender differences in ability. Although the gender differences in competitiveness we show in this paper may be task specific (as is true for any task used), our focus is on the comparison of the measures.

We conducted the main experiment in a university laboratory with a total of 210 participants in two treatments. The binary-measure treatment, which tests behavior on the extensive margin, consisted of 126 participants (71 women), with 6 participants per session. Seven sessions of this treatment were conducted in winter 2012, seven in fall 2013, and seven in fall 2014.[3]

The linear-measure treatment, testing behavior on the intensive margin, consisted of 14 experimental sessions with six participants in each session. Seven sessions were conducted in winter 2012 and seven in spring 2013. The linear-measure treatment consisted of 84 total participants (44 women). On average, participants across both treatments earned $8.20 including the show-up fee.[4] Online Appendix B reports all of the instructions.

2.1. Measures of Competitiveness

The only difference between the binary (extensive margin) and linear (intensive margin) measures of competitiveness was in the decisions participants made about their compensation for the task. Each participant made this choice before the beginning of the task.

2.1.1. Competitiveness on the Extensive Margin. The typical elicitation of competitiveness focuses on choices on the extensive margin. This measure entails a binary choice between two compensation schemes: a tournament compensation scheme (T) and a piece-rate compensation scheme (PR). The piece-rate scheme is based on individual performance alone: participants are paid $1 per successful toss independent of others' performances. The tournament compensation scheme pays $3 per successful toss if a participant wins against a randomly chosen opponent. Participants know that the opponent will be chosen ex post from the entire pool of participants from the same session—men and women, and not just those who chose to compete. To eliminate the effect of altruistic preferences on the decision to compete, the choice of whether to compete does not affect the success of the person with whom participants are matched; only their relative success in the task is relevant. The participant receives no payment if she loses the competition and $1 per successful toss in the case of a tie.

2.1.2. Competitiveness on the Intensive Margin. To measure the intensity of individuals' preferences for competitiveness, we introduce a measure that asks participants to choose a linear combination of tournament compensation and piece-rate compensation to compose their overall payoffs. That is, the decision maker receives 100 points and is asked to choose how much of it, t, she wishes to invest in the tournament option, T, and how much to invest in the piece-rate option, PR. At the end of the experiment, participants are paid \$1 for every 100 points earned. The payoffs are then $(100 - t + 3t) \times$ (the number of successful tosses) if the participant scores higher than her opponent and $(100 - t) \times$ (the number of successful tosses) if the participant scores lower than her opponent. In case of a tie, the participants simply get 100 times the number of successful tosses, or $(100 - t + t) \times$ (number of successful tosses). The total compensation for the task is calculated according to $\Pi = (t/100)\pi^T + (1 - t/100)\pi^{PR}$, with π^T as her tournament payoff and π^{PT} as her piece-rate payoff. This allocation to the tournament, t, is our measure of competitiveness. An individual is deemed more competitive than another if she chooses to include a greater amount of the tournament payoff in her chosen payoff combination than another individual.

2.1.3. Additional Measures. In addition to eliciting levels of competitiveness, in all sessions but the fall 2013 ones we measured other factors that may affect competitiveness. With these measures, we can observe the effect of hedging uncertainty and beliefs. There may, of course, be other factors contributing to gender differences in competitiveness in some settings, such as social preferences (see, e.g., Balafoutas et al. 2012). However, in our case we rule out this factor by design. The instructions are reported in Online Appendix C.

Confidence. We used two measures that we administered before the start of the ball-tossing task but after the choice of an incentive scheme. First, we asked participants to guess their expected number of successful tosses on a scale from 0 to 10 ("How many successful tosses do you think you will make?"). We label this measure *Expected performance*. Second, we asked participants to state the expected likelihood of winning against a random opponent, as a percentage from 0 to 100 ("What do you believe is the probability that you will make more successful tosses than a randomly selected opponent?"), which we refer to as *Confidence of winning*. Measuring beliefs is always a tricky task. We decided not to incentivize this belief elicitation in order to keep the instructions simple and avoid cross-influence between beliefs and effort on the task. Participants have no strategic reason to misreport their beliefs in our experiment.

Risk Attitudes. We elicited risk attitudes using two different measures. First, after making the choice but before performing the task, we elicited risk attitudes through the multiple price list measure of risk aversion (MPL; see Holt and Laury 2002). The measure was incentive compatible. We denote this measure as *Risk aversion*. At the end of the experiment, we compensated decisions made using this measure (see the payment procedure below). We also elicited self-assessed risk taking on a scale from 1 to 10 using the following question: "Please answer the following question using a 1–10 scale, where 1 = *completely unwilling* and 10 = *completely willing*: Rate your willingness to take risks in general." This measure is adapted from Dohmen et al. (2011), who find it to be predictive of risky behaviors and of participants' choices in an incentivized risk task.

Ambiguity Aversion. We also assessed ambiguity preferences with an MPL over known and unknown lotteries. As is typical of MPL, participants were presented with a series of 20 decisions. Each decision entailed a choice over a known and an unknown lottery. Similar to the risk measure, participants had to indicate a "switch point"—the point at which they decided to switch from choosing to be paid according to the known lottery to choosing to be paid according to the unknown lottery. This switch point serves as a measure of aversion to ambiguity and represents the premium the agent is willing to pay to avoid the ambiguous outcome. At the end of the experiment, we compensated responses for the two MPLs (risk and ambiguity measures). In particular, participants received payment for one of the two MPLs, determined at random by a coin flip.

2.2. Procedure

We invited participants to the lab using standard recruiting procedures. Each session had six participants. We invited more participants to ensure that we had six people per session, and if more than six people showed up, we dismissed the extra participants. Our goal was to have gender-balanced sessions, with three women and three men. However, because we could not recruit participants by gender, we could not always meet this goal. After being seated at their computer station, participants received the instructions. The instructions explained that participants would perform a ball-tossing task and that they had to decide how to be compensated for it. To make sure participants understood the instructions before they made their decision, an experimenter read the instructions out loud. The experimenter also gave a demonstration of the task without actually tossing the ball into the basket. Participants did not practice the task before making their compensation choice.

Throughout the experiment, gender was not made salient. Participants only knew that if they selected the tournament, their performance would be compared with the performance of a random opponent in the room. Most of the sessions were gender balanced. The computer stations in the lab faced the walls of the laboratory and were separated by dividers, preventing participants from looking at each other without completely turning.

In all sessions but the fall 2013 ones, after making their choices, participants filled out a short survey aimed at eliciting their confidence, as explained above. Next, they received two separate envelopes. One envelope contained the instructions and decision sheet for the risk-attitudes measure, and the other envelope contained the instructions and decision sheet for the ambiguity-preferences measure.[5] Participants were informed that in this portion of the experiment, they would be paid according to the realization of one decision across both tasks, and that the task determining their payment would be randomly selected by means of a coin toss at the end of the experiment.

We then directed one participant at a time to a separate room to perform the ball-tossing task while the rest of the participants waited at their computer stations. No communication was allowed between participants at any moment throughout the experiment. At the end of the experiment, participants filled out a short survey of basic demographic information, asking them for their age, ethnicity, spoken language, major, and GPA. The questionnaire also contained the self-reported measure of risk. At the end of the experiment, we paired each participant anonymously with a random opponent from the same session. We then paid them according to their choice of compensation-scheme offerings, their performance, and, in the relevant cases, the outcome of the tournament. We then also paid each participant for one random decision from either the risk or the ambiguity measure, depending on the outcome of the coin toss.

3. Results

3.1. Competitiveness on the Extensive Margin

Table 1 summarizes the tournament-entry results for the three rounds of experimental sessions. We observe no statistical difference across the three waves of sessions in the proportion of women and men who selected the tournament (men: $\chi^2(2) = 3.02$, $p = 0.22$; women: ($\chi^2(2) = 0.35$, $p = 0.84$). Therefore, we pool all of the binary-task data together for the analyses. Of the 126 individuals in the sample, 52.4% chose to participate in the tournament. Of the women, 32.4% chose to participate in the tournament, whereas 78.2% of men chose the tournament. This difference is statistically significant ($\chi^2(1) = 26.05$, $p < 0.001$). This gender gap in competitiveness replicates the results of previous research in selection into competitive environments using the same or a different task (see Online Appendix A2 for a review of the gender gap detected in previous work).

We further investigate this result using a regression framework. Table 2 reports the results of different specifications of a probit model in which we regress a tournament-entry dummy variable on a female dummy. As the table shows, women are significantly less likely to choose the competitive scheme. The estimated marginal effect reported in column (1) suggests women are 36.4 percentage points less likely to enter the tournament than men. Because the women-to-men ratio was not constant across sessions, in column (2), we control for the gender composition of the sessions. In particular, whereas most of the observations (66.7%) come from gender-balanced sessions (three men and three women), some of the sessions (28.6% of the observations) were characterized by a majority of women (three sessions had four women and two men, and three sessions had five women and one man), and one session (4.8% of the observations) had a majority of men. Because the gender of a potential competitor may affect participants' willingness to compete, participants' choices might have differed in the unbalanced sessions. We control for this heterogeneity in the gender composition of the sessions by adding to the model a variable indicating the women-to-men ratio in each session.[6] When adding this variable to the model, we find women to be 39.9 percentage points less likely to select into the tournament (column (2a)). Furthermore, running a regression on only the 66.7% of the participants ($N = 84$) who took part in gender-balanced sessions leads to a similar result, with an estimated gender gap of 36.6 percentage points ($p < 0.001$). Online Appendix A4 reports the results of this regression and of other specifications with controls.

When considering participants' performance conditional on tournament choice, we find subjects completed an average of 2.48 tosses (SD = 1.68). Furthermore, we find men performed better than women under

Table 1. Summary Statistics—Competitiveness on the Extensive Margin

	Fraction of participants who selected the tournament										
	2012 sessions ($N = 42$)		2013 sessions ($N = 42$)		2014 sessions ($N = 42$)		Pooled data ($N = 126$)				
	Fraction	N	Fraction	N	Fraction	N	Fraction	N	Min	Median	Max
Men	0.71	14	0.70	20	0.90	21	0.78	55	0	1	1
Women	0.29	28	0.36	22	0.33	21	0.32	71	0	0	1
All	0.43	42	0.52	42	0.62	42	0.52	126	0	1	1

Table 2. Probit Regression of Tournament-Entry Decision

	(1)	(2a)	(2b)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	−0.364***	−0.399***	−0.422***	−0.423***	−0.334***	−0.318***	−0.313***	−340***	−0.373***	−0.359***
	(0.046)	(0.086)	(0.104)	(0.103)	(0.112)	(0.114)	(0.116)	(0.123)	(0.116)	(0.112)
Gender composition		−0.001	−0.001	−0.001	−0.016	−0.021	−0.016	−0.004	−0.002	−0.014
		(0.036)	(0.036)	(0.037)	(0.048)	(0.047)	(0.049)	(0.049)	(0.046)	(0.044)
Expected performance				0.005						
				(0.027)						
Confidence of winning					0.011**	0.011**	0.011**	0.011**	0.010**	0.010**
					(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Self-reported risk						−0.001	−0.001	−0.000	0.016	0.014
						(0.034)	(0.033)	(0.034)	(0.034)	(0.034)
Risk aversion							0.046	0.049	0.051	0.056*
							(0.030)	(0.031)	(0.032)	(0.033)
Ambiguity aversion								−0.019	−0.021*	−0.020*
								(0.012)	(0.012)	(0.012)
Age									−0.005	−0.005
									(0.026)	(0.026)
Asian									−0.184	−0.149
									(0.128)	(0.121)
Nonnative speaker									0.044	0.027
									(0.179)	(0.170)
GPA										−0.091
										(0.121)
Year dummy	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Major dummies	N	N	N	N	N	N	N	N	Y	Y
Observations	126	126	84	84	84	83	80	80	79	77
Pseudo R^2	0.156	0.166	0.216	0.217	0.307	0.304	0.312	0.329	0.389	0.384

Notes. The table presents marginal effects estimated from probit regression. The dependent variable is *Choice of tournament* (equal to 1 for tournament and 0 for piece-rate). *Gender composition* refers to the women-to-men ratio in each session. *Expected performance* refers to the estimated number of successful tosses. *Confidence of winning* refers to subjects' expected likelihood of winning against a random opponent from the same session. *Self-reported risk* refers to the self-reported willingness to take risks. *Risk aversion* refers to the incentivized measure of risk aversion. *Ambiguity aversion* refers to the incentivized measure of ambiguity aversion. *Age* refers to subjects' age. *Asian* is a dummy variable coded as 1 if subjects were of Asian ethnicity and 0 otherwise. *Nonnative speaker* is a dummy variable coded as 1 if subjects were not English native and 0 otherwise. *GPA* refers to the demeaned self-reported GPA. Major dummies are dummy variables for the following majors: engineering and math, social science, and the residual majors (which include literature, art, communication, and undeclared), with science as the baseline. Marginal effects are evaluated at a man in a 2012 gender-balanced session and at the mean for all of the other variables. Robust standard errors are in parentheses.

***$p < 0.01$; **$p < 0.05$; *$p < 0.10$.

the tournament ($z = 2.277$, $p = 0.02$; Mann–Whitney), which is in line with previous literature (Gneezy et al. 2003), and marginally better than women under the piece rate ($z = 1.717$, $p = 0.09$; Mann–Whitney).[7] Both results become insignificant if we exclude the top participants who perform five or more successful tosses (11.66% of the participants; $z = 1.438$, $p = 0.15$ for the tournament; $z = 1.289$, $p = 0.20$ for the piece rate; Mann–Whitney). If we limit our regression analysis reported in Table 2 to the sample of participants with no differences in performance, we still observe that females are 35.5 percentage points less likely than men to select the tournament ($p < 0.001$, $N = 106$; see Online Appendix A5).

3.1.1. Determinants of Competitiveness on the Extensive Margin.

In this section, we investigate whether the gender gap in tournament entry is driven by gender differences in participants' confidence of winning the competition, in risk preferences, and in ambiguity aversion. As a reference, in column (2b), we report the same analysis of column (2a) on the restricted sample of participants ($N = 84$) from whom we have data on such measures.

Confidence. Both men and women were overconfident regarding the number of successful tosses they would perform in the task. On average, participants expected to successfully make 5.04 tosses (SD = 1.93, $N = 84$), which is more than the actual average number of tosses these participants completed ($\text{Mean}_{\text{tosses}} = 2.48$, $z = 6.99$, $p < 0.001$; Wilcoxon test). In line with previous literature, we find a gender gap in confidence. In particular, men expected an average of 5.79 tosses (SD = 1.75, $N = 35$), whereas women expected to successfully complete an average of 4.51 tosses (SD = 1.90, $N = 49$), with the two distributions being significantly different ($z = 2.550$, $p = 0.01$; Mann–Whitney). When we look at participants' confidence of winning against a random opponent, we find a similar gender difference. On average, men's reported expected likelihood of winning is 63.13%, whereas women's is 43.29% ($z = 4.731$, $p < 0.001$; Mann–Whitney). The two measures of confidence are strongly correlated ($r = 0.62$, $p = 0.001$).

We add these variables to the regression model reported in Table 2. Because the two variables are highly correlated, we do not add both of them to the same model. Column (3) shows that participants' expected performance is not correlated with tournament entry. Adding this variable to the model does not reduce the gender gap in competitiveness. Column (4) instead shows that participants' confidence in winning against a random opponent is significantly correlated with the tournament-entry decision. Adding this variable to the model reported in column (2) reduces the gender gap by about nine percentage points. This result is consistent with the results observed in the literature (e.g., Niederle and Vesterlund 2007). Although part of the gender gap in tournament entry can be attributed to confidence about the likelihood of winning, a substantial gap between men's and women's choices to compete remains.

Risk. The two measures are not significantly correlated ($r = -0.134$, $p = 0.238$). We do not observe gender differences in the incentivized risk-aversion measure (switch-point $\text{Mean}_{\text{men}} = 6.34$, SD = 1.81 versus $\text{Mean}_{\text{Women}} = 6.78$, SD = 1.99, $z = -1.022$, $p = 0.307$; Mann–Whitney). However, we do find a gender difference in the self-reported risk measure, with men being less risk averse than women ($\text{Mean}_{\text{Men}} = 6.77$, SD = 1.80 versus $\text{Mean}_{\text{Women}} = 5.48$, SD = 1.87, $z = -3.27$, $p = 0.001$; Mann–Whitney). When adding the risk measures to the model (columns (5) and (6)), we find no robust correlation between the risk measures and the choice of tournament, and the gender gap does not substantially change.

Ambiguity. We do not observe gender differences in ambiguity aversion. The average switching point of men was 11.86 (SD = 4.83), whereas that of women was 11.62 (SD = 4.73, $z = 0.26$, $p = 0.80$; Mann–Whitney). Adding this variable to the regression model shows no significant correlation with the tournament-entry decision, though we detect a marginally significant negative correlation when adding demographic controls to the model. Importantly, controlling for ambiguity preferences in addition to risk preferences and confidence does not substantially change the gender gap.

Demographics Controls. In columns (8) and (9), we control for additional demographic variables we collected at the end of the experiment. Online Appendix A3 reports a detailed description of the demographic composition of our sample.

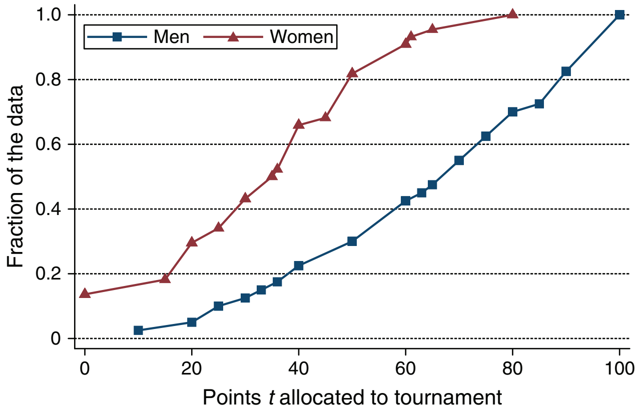
3.2. Competitiveness on the Intensive Margin

Table 3 presents the summary of tournament allocations by gender. We find no difference between the distributions of points allocated in the two rounds of sessions ($z = -0.987$, $p = 0.32$; Mann–Whitney). Hence, we pool the data for the analyses. Overall, participants allocated an average of 50.11 points (SD = 28.09) to the tournament. The median allocation was 50 points. The average number of points allocated to the tournament is markedly different for men than for women: on average, women allocated 35.27 points to the tournament (SD = 21.19), whereas men allocated 66.43 points to the tournament (SD = 25.74). The distribution of points allocated to the tournament differs by gender ($z = -4.99$, $p < 0.001$; Mann–Whitney). Importantly, the shapes of the distributions are also different. This fact is evident in the empirical cumulative distribution function (CDF) in Figure 2 and in the smoothed probability density function (PDF) in Figure 3. The distribution for men is visibly shifted to the right, along the axis of competitiveness, with respect to the distribution for women. The summary statistics provided in Table 3 also depict this difference in distributions; the quartiles calculated for each population are strikingly different. For example, only the most competitive 25% of women allocated 50 points or more to the tournament, whereas only the 25% least competitive men allocated fewer than 50 points. Only the most competitive

Table 3. Summary Statistics —Competitiveness on the Intensive Margin

	Tournament allocations							
	2012	2013	Pooled					
	Mean (SD)	Mean (SD)	Mean (SD)	Min	25th percentile	Median	75th percentile	Max
Men	63.8 (27.86)	69.05 (23.86)	66.43 (25.74)	10	50	70	90	100
Women	31.64 (20.39)	38.91 (21.81)	35.27 (21.19)	0	20	35.5	50	80
All	46.95 (28.93)	53.26 (28.81)	50.11 (28.09)	0	30	50	70	100

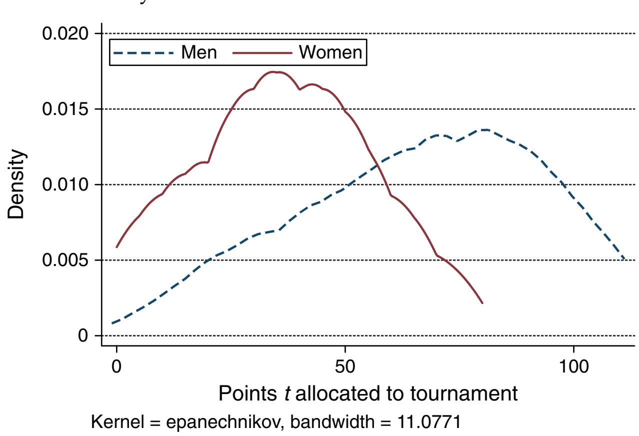
Figure 2. (Color online) Empirical CDF of Tournament Allocations t by Gender



woman, who allocated 80 points, allocated more points than the median man (who allocated 70 points).

Table 4 reports the empirical results from different specifications of anordinary least squares (OLS) regression where the number of points t allocated to the tournament option is regressed on a gender dummy. The first specification of the model reported in column (1) shows that women allocated significantly fewer points to the tournament scheme than men ($\beta = -31.15$, $p < 0.001$). In the experiment, all but two sessions were gender balanced (three men and three women). In the two unbalanced sessions, the fraction of women to men was four to two. When controlling for the gender composition of the session by adding a variable indicating the women-to-men ratio in the sessions to the regression model (column (2)), we find that women still allocated significantly fewer points to the tournament option. The coefficient of the gender-composition variable indicates that participants allocated more points to the tournament in the sessions with a higher fraction of women, though one should exercise caution in interpreting this result, because it is based on only two

Figure 3. (Color online) Smoothed PDF of Tournament Allocations by Gender



sessions with more women than men. In addition, we find a nonsignificant effect of the interaction between gender and gender composition, suggesting gender composition does not affect women and men differently ($\beta = 2.64$, $p = 0.852$); hence, we do not include the interaction term in the models in Table 4.

When looking at participants' performance after their allocation decision, we find that, on average, participants successfully completed 1.90 tosses (SD = 1.65). Men performed marginally better than women ($z = 1.85$, $p = 0.064$), though this difference becomes insignificant if we exclude participants who performed more than five tosses (5.95% of the participants; $z = 1.335$, $p = 0.182$). If we limit the regression analysis reported above to the sample of participants with no differences in performance, we still find that women allocated 29.78 fewer points to the tournament than men ($p < 0.001$, $N = 79$; see Online Appendix A5).

3.3. Determinants of Gender Differences in Competitiveness on the Intensive Margin

Confidence. As above, participants were overconfident about their task ability. On average, participants expected to successfully complete 4.90 tosses (SD = 1.90), which is more than the average number of tosses these participants actually completed ($\text{Mean}_{\text{tosses}} = 1.90$, $z = 7.94$, $p < 0.001$; signed-rank test). However, men were significantly more confident about their performance than women ($\text{Mean}_{\text{Men}} = 5.75$, SD = 1.71, $N = 40$ versus $\text{Mean}_{\text{Women}} = 4.28$, SD = 1.89, $N = 44$), with the two distributions being statistically different ($z = 3.35$, $p < 0.001$; Mann–Whitney). Similarly, men indicated that their chance of winning against a random opponent was 65.15% (SD = 16.19), whereas women reported a chance of winning of 43.70% (SD = 17.55, $z = 5.05$, $p < 0.001$; Mann–Whitney). The two confidence measures are strongly correlated ($r = 0.53$, $p < 0.001$).

We add these variables to the regression model reported in Table 2. Column (3) shows that individuals' expected performance marginally correlates with the number of points they allocated to the tournament. However, adding expected performance to the model does not reduce the gap between men's and women's average tournament allocations. Column (4) shows that one's confidence about winning is positively correlated with tournament allocations. That is, more confident individuals allocate more points to the tournament, regardless of their gender. However, considering a man and a woman with the same level of confidence, a woman allocates fewer points to the tournament. This result shows that although accounting for differences in confidence between men and women substantially reduces the gap in tournament allocations, confidence about winning cannot account for it entirely.

Table 4. OLS Regression of Tournament Allocation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	−31.15***	−32.30***	−28.80***	−17.64***	−14.68**	−12.58**	−16.32***	−19.21**	−16.93**
	(5.17)	(5.17)	(5.42)	(5.86)	(6.00)	(5.72)	(5.82)	(7.32)	(7.71)
Gender composition		14.01**	12.86*	8.79**	4.72	4.84	4.85	3.07	3.92
		(6.35)	(7.01)	(4.29)	(4.60)	(4.40)	(4.24)	(4.58)	(4.65)
Expected performance			2.64*						
			(1.44)						
Confidence of winning				0.664***	0.533***	0.594***	0.539***	0.511***	0.511***
				(0.118)	(0.137)	(0.135)	(0.135)	(0.141)	(0.144)
Self-reported risk					3.22**	3.16**	3.03*	3.29*	2.82
					(1.51)	(1.52)	(1.64)	(1.77)	(1.86)
Risk aversion						0.533	0.567	0.373	−0.102
						(1.09)	(1.08)	(1.19)	(1.21)
Ambiguity aversion							−0.809	−0.907	−0.743
							(0.549)	(0.571)	(0.598)
Age								1.41	1.63
								(1.40)	(1.35)
Asian								−5.22	−5.69
								(5.34)	(5.08)
Nonnative speaker								3.33	2.13
								(5.44)	(5.74)
GPA									−10.24
									(6.24)
Constant	66.43***	47.86***	33.91***	10.96	1.98	−6.21	8.40	15.32	18.46
	(4.06)	(8.66)	(12.69)	(10.84)	(12.21)	(16.93)	(19.85)	(20.43)	(21.26)
Year dummy	N	Y	Y	Y	Y	Y	Y	Y	Y
Major dummies	N	N	N	N	N	N	N	Y	Y
Observations	84	84	84	84	82	78	77	75	73
R^2	0.311	0.350	0.382	0.507	0.538	0.573	0.587	0.601	0.610

Notes. The table reports OLS estimates. The dependent variable is *Points allocated to tournament*. *Gender composition* refers to the women-to-men ratio in the session. *Expected performance* refers to the estimated number of successful tosses. *Confidence of winning* refers to subjects' expected likelihood of winning against a random opponent from the same session. *Self-reported risk* refers to the self-reported willingness to take risks. *Risk aversion* refers to the incentivized measure of risk aversion. *Ambiguity aversion* refers to the incentivized measure of ambiguity aversion. *Age* refers to subjects' age. *Asian* is a dummy variable coded as 1 for subjects of Asian ethnicity. Nonnative speaker is a dummy variable coded as 1 if subjects were not English natives and 0 otherwise. *GPA* refers to the demeaned self-reported GPA. Major dummies are dummy variables for the following majors: engineering and math, social science, and the residual majors (which include literature, art, communication, and undeclared), with science as the baseline. Robust standard errors are reported in parentheses.

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Risk. As in the binary-measure sessions, we find a significant gender difference in the self-reported risk measure. On average, men rated themselves as more willing to take risks than women ($\text{Mean}_{\text{Men}} = 6.79$, SD = 1.70 versus $\text{Mean}_{\text{Women}} = 4.98$, SD = 1.79); the two distributions are different ($z = 4.18$, $p < 0.001$; Mann–Whitney). We do not observe a gender difference in the incentivized measure of risk (men's average switch point = 7.24, SD = 1.91; women's average switch point = 7.22, SD = 2.04 ($z = -0.05$, $p = 0.964$; Mann-Whitney)). The two risk measures are not correlated ($r = -0.023$, $p = 0.84$).

In columns (5) and (6) of Table 4, we explore whether risk preferences affect participants' tournament-allocation decisions by adding these variables to the regression reported in column (4). Column (5) shows that the self-reported measure of risk partly accounts for participants' tournament allocation. We observe that participants who described themselves as more likely to take risks allocated more points to the tournament. However, the gender differences in tournament allocation remains. In column (6), we show that the result is robust to adding the incentivized measure of risk to the model, which is not correlated with tournament allocation.

Ambiguity. We find a significant gender difference in ambiguity aversion in this sample, with men being more averse to ambiguity than women (men: average switching point = 12.63, SD = 4.99; women: 9.65, SD = 3.74); the two distributions are significantly different ($z = 2.69$, $p = 0.007$; Mann–Whitney). When adding the ambiguity measure to the regression model, we find that it does not contribute to explaining tournament-allocation decisions. Column (7) shows that when we control for this variable, women still allocate significantly fewer points to the tournament.

Demographic Controls. In columns (8) and (9), we control for additional demographic variables we collected at the end of the experiment (see Online Appendix A3 for a detailed description of such variables).

Overall, these results show that confidence about winning, as well as risk preferences, can partly account for participants' allocation decisions. However, when accounting for beliefs and risk preferences, women still allocate significantly fewer points to the tournament.

Confidence and Risk in the Two Measures of Competitiveness. Our results also show that, compared with the extensive-margin measure of competitiveness, measuring competitiveness through a linear-allocation task provides a finer characterization of the relationship between competitiveness and other economic preferences. In the binary elicitation, our analyses show that participants' confidence significantly correlates with tournament entry, even when controlling for gender. With respect to risk, whereas the self-reported measure of willingness to take risk correlates with the tournament-entry decisions in a probit regression of risk on tournament entry and no other control variables by increasing subjects' likelihood of entering the tournament by eight percentage points ($p = 0.006$), the effect becomes nonsignificant when we control for confidence ($p = 0.363$) or gender ($p < 0.175$). In our sample, measuring competitiveness with a binary choice does not allow us to identify a relationship between risk preferences and competitiveness. By contrast, the finer measure of competitiveness that we introduce in this paper not only captures the relationship between competitiveness, confidence, and gender but also allows us to identify a relationship between (self-reported) risk preferences and competitiveness.

3.4. The Gender Gap in Competitiveness in the Two Measures

In this section, we examine the women-to-men ratio across the different percentiles of the distribution of competitiveness of the linear (intensive margin) measure and compare it to the one captured by the binary (extensive margin) measure. For the binary measure, we divide the fraction of women who choose the tournament out of all women in the treatment, $w(T)$, by the fraction of men who choose the tournament out of all men, $m(T)$. For the linear measure, we calculate this ratio at various percentiles of the distribution of competitiveness, represented by the points t allocated to the tournament. At the different percentiles in the distribution, we consider $w(t)$ the fraction of women who allocate t points or higher to the tournament option and $m(t)$ the fraction of men who allocate t points or higher to the tournament option, and we calculate the women-to-men ratio, $w(t)/m(t)$.

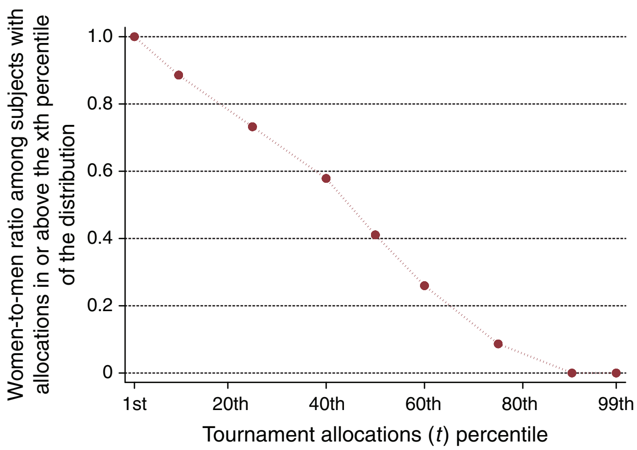
Our results, reported in Table 5, reveal that the binary measure captures a women-to-men ratio of 0.41 to 1, with about one-third of the women and a little over two-thirds of the men choosing to compete. Is this ratio constant in all of the percentiles above a certain cutoff of the distribution of competitiveness in the linear measure? Our analysis reveals that it is not; we find that this ratio decreases for increasing degrees of competitiveness. Figure 4 provides a graphical representation of the decay in the women-to-men ratio at the higher percentiles of the distribution of competitiveness. This ratio is plotted as a function of the number of points t allocated to the tournament; these points are represented on the x axis by percentile ranks of the entire sample of men and women. The first percentile of the distribution serves as a benchmark where the women-to-men ratio is 1, because it is computed on all participants who allocated zero or more points to the tournament. If men and women were equally distributed across the distribution of competitiveness, we should observe this ratio at all percentiles of the distribution. When focusing on the top 50% of the distribution, the women-to-men ratio is 0.41 to 1, which is similar to the ratio we observe in the binary measure. However, the gap between the fractions of men and women widens substantially when moving toward the upper tail of the distribution. The women-to-men ratio becomes 0.09 on the 75th percentile and above;

Table 5. Women-to-Men Ratio in the Two Measures

Binary		Linear						
		Percentile						
	Tournament choice	1th	10th	25th	50th	75th	90th	95th
t to tournament	—	$t \geq 0$	$t \geq 15$	$t \geq 30$	$t \geq 50$	$t \geq 70$	$t \geq 90$	$t = 100$
Fraction of men	0.78	1	0.98	0.90	0.78	0.53	0.28	0.18
Fraction of women	0.32	1	0.86	0.66	0.32	0.05	0	0
Women-to-men ratio	0.41	1	0.88	0.73	0.41	0.09	0	0

Notes. Fraction of men refers to the fraction of men who chose a tournament allocation that is equal to or greater than t out of all men in the treatment. *Fraction of women* refers to the fraction of women who chose a tournament allocation that is equal to or greater than t out of all women in the treatment. *Women-to-men ratio* refers to fraction of women divided by fraction of men.

Figure 4. (Color online) Women-to-Men Ratio Along the Distribution of Competitiveness (Intensive-Margin Measure)



it becomes 0 at the 90th percentile, because all of the participants in the top 10% of the distribution are men.

We investigate whether the estimated gender gap at the median and the top percentiles of the distribution of competitiveness differ from the estimated gender gap observed with the binary measure. To compare the observations from the two elicitations of competitiveness, we consider the top x percentile of all players (men and women) ranked in terms of the linear measure of competitiveness. We classify as "competitive" the participants in the top x percentile and as "noncompetitive" the participants who are below the top x percentile. In the binary measure, we code as "competitive" the participants who chose the tournament and as "noncompetitive" the participants who chose the piece rate. We compare whether the proportion of men and women among the competitive participants are the same across the two measures. A Fisher exact test reveals that in the top 50th percentile of the distribution of competitiveness, the proportion of men and women does not differ across the two elicitations (p = 0.84; Fisher exact, two-sided). Indeed, the women-to-men ratio is similar in the two cases. This result confirms that the gender gap observed at the median of the distribution of competitiveness in the linear measure is similar to the gap we observe in the binary measure. However, when we compare the fraction of men and women at the 75th percentile of the distribution of competitiveness and above with the proportion of people who choose the tournament in the case of a binary choice, the two proportions differ ($p = 0.016$; Fisher exact, two-sided). Similarly, the proportion of men and women in the top 10th percentile of the distribution is significantly different from this proportion in the binary measure (p = 0.028; Fisher exact, two-sided). In Online Appendix A6, we repeat the analysis on participants from gender-balanced sessions only and find consistent results. In Online Appendix A6, we also report additional analyses exploring whether confidence and other risk preferences can account for whether participants are in the top (bottom) tail of the distribution of competitiveness, showing that confidence and risk correlate with whether participants are among the top 25% of the distribution but cannot account for all of the gender gap; conversely, we find that confidence alone accounts for whether individuals are in the bottom 25% of the distribution.

Taken together, our results show that although about one-third of the women show some level of competitiveness, the most competitive people in our sample are primarily men. The fraction of women among the most competitive participants is smaller than what is captured by a measure that relies on a binary choice. If those who are hired for very competitive jobs are drawn from the pool of individuals in the very upper tail of the distribution of competitiveness, the large gender gap we observe in the real world could also be partly because women are largely underrepresented at the top of the distribution. This gap does not need to appear when a person is considered for the job of a CEO; it could start at a much earlier stage in which the person is considering a future career path.

The results of the linear measure suggest that the distribution of competitive spirit is gender dependent. By comparing these results with the tournament-entry decision in the binary measure, we can shed further light on whether the minimum competitiveness levels required to enter a competition differ between men and women. Of the women in the extensive-margin treatment, approximately 32% chose to enter the tournament. Of the women in the intensive-margin treatment, approximately 32% allocated more than 45 points to the tournament. If we were to extend the results of the extensive-margin measure into the context of the intensive-margin measure, the "cutoff point," or decision rule for entering a competition, would be at about 45 points for women: women who allocate 45 points or more in the linear measure are the type that we assume chose to enter the tournament in the binary measure. By contrast, about 78% of men chose to enter the tournament in our binary-measure sample, whereas 77.5% of men allocated more than 40 points to the tournament. Thus, following the same reasoning as for women, the average cutoff point of men is about the same as that of women: a man allocating 40 points or fewer to the tournament would likely not choose to enter the tournament. These data suggest men and women use similar cutoffs in deciding whether to enter the tournament. Although this conclusion is based on the assumption that behavior is directly comparable between the two measures, the results of our experiments suggest that the gender difference in tournament entry is mostly due to differences in the shape of the distribution of competitiveness preferences rather than to gender differences in the minimum level of competitive spirit needed to enter a tournament.

To sum, by comparing the results obtained from the new measure with those of the binary measure, we can conclude that the linear measure of competitiveness provides a more accurate method for measuring the intensity of men's and women's preferences for competitiveness and for estimating the size of the gender gap at the top of the distribution. Furthermore, it allows for a deeper investigation of the relationship between competitiveness and other economic preferences. Finally, because it allows for a finer representation, the finer measure has the methodological advantage of, for a given power level, needing fewer observations to correctly reject a null hypothesis.[8]

4. Discussion

Competiveness is a personal preference with important economic implications; hence, understanding what affects the tendency to compete is useful for economic analysis. We propose a new elicitation procedure that allows us to obtain a refined measure of individual competitiveness. Previous measures did not allow for variation in levels of competitiveness and therefore masked the real size of the gender gap at the top of the distribution. As we argued in the introduction, elements that could go into competitiveness traits include the reaction to competitive incentives as well as the selection into competitive environments. Future research can use our findings in addition to other findings in the literature in a search for a more comprehensive measure.

Our results suggest that although women might choose to select into environments characterized by a moderate degree of competitiveness, they might opt out of highly competitive environments. Interestingly, we observe that 78% of men and 32% of women entered the tournament in the extensive-margin treatment, whereas men allocated 66% and women 35% of their points toward the tournament. The order of magnitude for both the intensive and extensive margins appears similar. Importantly, though, the new measure shows that, conditional on the decision to compete, men present stronger preferences for competitiveness. The gender differences in the very top of the competitiveness distribution are not captured by the binary measure but only by the new measure; the striking difference there is important in understanding the extent of gender differences. If successful careers in certain segments of the labor market demand a high level of competitiveness, we can reasonably project that a weaker preference for competition will lead fewer women to commit to such career paths. Our results suggest that women with highly competitive preferences may be the exception rather than the rule. These results have implications beyond career choice and financial success. An individual's competitiveness may also affect her likelihood of engaging in other competitive interactions, such as bargaining. Hence, attitudes toward competition may affect, for example, entry into wage negotiations, which in turn could have bearing on the wage gap between men and women in similar occupational positions (Babcock and Laschever 2003, Rigdon 2012).

Understanding the magnitude of the gender gap in competitiveness can also be helpful for developing different incentives schemes for men and women. A nice example is the recent paper by Petrie and Segal (2015), who test whether a price mechanism could be used to achieve gender balance. They vary the return on competition and find that if the rewards to competition are sufficiently large, women are willing to compete as much as men. They characterize people according to the minimum prize at which they choose to enter a tournament, showing that women choose to enter at significantly higher minimum prizes and that only a small fraction of the initial gender gap can be attributed to performance, beliefs, and general factors such as risk and feedback aversion (see also Ifcher and Zarghamee 2016).

Future research can use our setup to further the understanding of the underlying causes of the gender gap in labor markets, and the ways to model it, by using a design in which the same person is performing the same task under different incentives as in Niederle and Vesterlund (2007). See Charness et al. (2012) for a discussion of advantages and disadvantages of such an approach. Applications of this method could further investigate how other preferences affect gender differences in competitiveness observed in our experiment, both directly and indirectly.

Acknowledgments

The authors thank Roel van Veldhuizen, Anastasia Danilov, the department editor, the associate editor, three anonymous referees, and numerous conference participants for invaluable feedback and comments.

Endnotes

[1] For discussions regarding the estimated size of the gender gap in, for example, wages, and for possible economic explanations, see Altonji and Blank (1999), Bertrand and Hallock (2001), Bertrand (2011), and Goldin et al. (2006). For recent statistics, see Fairchild (2014).

[2] For further evidence on gender differences in selecting into competition, see Andersen et al. (2013), Balafoutas and Sutter (2012), Booth and Nolen (2012), Cason et al. (2010), Datta et al. (2013), Dohmen and Falk (2011), Ertac and Szentes (2011), Gneezy and Rustichini (2004), Gneezy et al. (2009), Healy and Pate (2011), Niederle et al. (2012), Sutter and Ruützler (2014), Vandegrift and Yavas (2009), Wozniak et al. (2014), Almås et al. (2015), Buser et al. (2014, 2016), Dargnies (2012), Flory et al. (2015), and Niederle and Vesterlund (2010, 2011).

[3] We originally had 84 participants per treatment but did not have data on confidence, risk, or ambiguity for half of the subjects who performed the binary task. To make use of that data and make sensible comparisons between treatments, we collected an additional 42 observations in the binary measure. We thank an anonymous referee for this suggestion.

[4] We found no difference in average earnings across the two measures ($8.46 in the binary elicitation versus $7.90 in the linear elicitation, t-test, $p = 0.36$). On average, men earned more than women in both measures (average earnings: $9.91 versus $7.27 in the binary elicitation, $p = 0.001$ t-test; $9.06 versus 6.84 in the linear elicitation, $p = 0.01$ t-test).

[5] Participants in the fall 2013 sessions did not complete the measures of confidence, risk and ambiguity attitudes, or grade point average (GPA). We collected additional data that include these measures in fall 2014. We thank an anonymous referee for the suggestion.

[6] Alternatively, considering (a) the total fraction of women in the session, (b) adding to the model dummy variables for sessions with more women than men and for sessions with more men than women, or (c) adding session dummies to the regression results in a similar gender gap.

[7] The data regarding the number of successful tosses are missing for one of the sessions ($N = 6$); they were not collected as a result of a mistake in the experimental procedures.

[8] A simple power calculation shows that for the extensive-margin measure, detecting an effect size analogous to the one observed in our experiment (32.4% of women as opposed to 78.2% of men selecting the tournament) with 90% power at a two-sided 5% significance level using a chi-square test would require a sample size of 46 subjects. For the intensive-margin measure, detecting an effect analogous to the one observed in our study (tournament allocations: $\text{Mean}_{\text{men}} = 66.43$, $\text{SD}_{\text{men}} = 25.74$ versus $\text{Mean}_{\text{women}} = 35.27$, $\text{SD}_{\text{women}} = 21.19$) with 90% power at a two-sided 5% significance level using a t-test requires a sample size of only 28 participants.

References

Almås I, Cappelen AW, Salvanes KG, Sørensen EØ, Tungodden B (2015) Willingness to compete: Family matters. *Management Sci.* 68(8):2149–2162.

Altonji JG, Blank RM (1999) Race and gender in the labor market. Ashenfelter OC, Card D, eds. *Handbook of Labor Economics* Vol. 3C, (Elsevier, Amsterdam), 3143–3259.

Andersen S, Ertac S, Gneezy U, List JA, Maximiano S (2013) Gender, competitiveness, and socialization at a young age: Evidence from a matrilineal and a patriarchal society. *Rev. Econom. Statist.* 95(4):1438–1443.

Babcock L, Laschever S (2003) *Women Don't Ask: Negotiation and the Gender Divide* (Princeton University Press, Princeton, NJ).

Balafoutas L, Sutter M (2012) Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science* 335(6068):579–582.

Balafoutas L, Kerschbamer R, Sutter M (2012) Distributional preferences and competitive behavior. *J. Econom. Behav. Organ.* 83(1): 125–135.

Bateman AJ (1948) Intra-sexual selection in *Drosophila*. *Heredity* 2(Part 3):349–368.

Bertrand M (2011) New perspectives on gender. Ashenfelter O, Card D, eds. *Handbook of Labor Economics*, Vol. 4B, (Elsevier, Amsterdam), 1543–1590.

Bertrand M, Hallock KF (2001) The gender gap in top corporate jobs. *Indust. Labor Relations Rev.* 55(1):3–21.

Booth A, Nolen P (2012) Choosing to compete: How different are girls and boys? *J. Econom. Behav. Organ.* 81(2):542–555.

Buser T, Dreber A, Mollerstrom J (2016) The impact of stress on tournament entry. *J. Experiment. Econom.*, ePub ahead of print November 11, doi:10.1007/s10683-016-9496-x.

Buser T, Niederle M, Oosterbeek H (2014) Gender, competitiveness and career choices. *Quart. J. Econom.* 129(3):1409–1447.

Cason TN, Masters WA, Sheremeta RM (2010) Entry into winner-take-all and proportional-prize contests: An experimental study. *J. Public Econom.* 94(9–10):604–611.

Charness G, Gneezy U (2012) Strong evidence for gender differences in risk taking. *J. Econom. Behav. Organ.* 83(1):50–58.

Charness G, Gneezy U, Kuhn M (2012) Experimental method: Between-subject and within-subject design. *J. Econom. Behav. Organ.* 81(1):1–8.

Croson R, Gneezy U (2009) Gender differences in preferences. *J. Econom. Literature* 47(2):448–474.

Dargnies M-P (2012) Men too sometimes shy away from competition: The case of team competition. *Management Sci.* 58(1):1982–2000.

Darwin C (1871) *The Descent of Man, and Selection in Relation to Sex* (John Murray, London).

Datta GN, Poulsen A, Villeval MC (2013) Gender matching and competitiveness: Experimental evidence. *Econom. Inquiry* 51(1): 816–835.

Dekel E, Scotchmer S (1999) On the evolution of attitudes towards risk in winner-take-all games. *J. Econom. Theory* 87(1):125–143.

Dohmen T, Falk A (2011) Performance pay and multidimensional sorting: Productivity, preferences, and gender. *Amer. Econom. Rev.* 101(2):556–590.

Dohmen T, Falk A, Huffman D, Sunde U, Schupp J, Wagner GG (2011) Individual risk attitudes: Measurement, determinants, and behavior consequences. *J. Eur. Econom. Assoc.* 99(3): 522–550.

Eckel CC, Grossman PJ (2008) Men, women, and risk aversion: Experimental evidence. Plott CR, Smith VL, eds. *Handbook of Experimental Economics Results* Vol. 1, (North-Holland, Amsterdam), 1061–1073.

Ertac S, Szentes B (2011) The effect of performance feedback on gender differences in competitiveness: Experimental evidence. Working paper, Koç University, Istanbul.

Fairchild C (2014) Number of Fortune 500 women CEOs reaches historic high. *Fortune* (June 3), http://fortune.com/2014/06/03/number-of-fortune-500-women-ceos-reaches-historic-high/.

Flory JA, Leibbrandt A, List JA (2015) Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions. *Rev. Econom. Stud.* 82(1):122–155.

Gneezy U, Potters J (1997) An experiment on risk taking and evaluation periods. *Quart. J. Econom.* 112(2):631–645.

Gneezy U, Rustichini A (2004) Gender and competition at a young age. *Amer. Econom. Rev.* 94(2):377–381.

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

Gneezy U, Niederle M, Rustichini A (2003) Performance in competitive environments: Gender differences. *Quart. J. Econom.* 118(3): 1049–1074.

Goldin C, Katz LF, Kuziemko I (2006) The homecoming of American college women: The reversal of the college gender gap. *J. Econom. Perspect.* 20(4):133–156.

Günther C, Ekinci NA, Schwieren C, Strobel M (2010) Women can't jump? An experiment on competitive attitudes and stereotype threat. *J. Econom. Behav. Organ.* 75(3):395–401.

Hausmann R, Tyson LD, Zahidi S (2010) The global gender gap report 2009. Report, World Economic Forum, Geneva.

Healy A, Pate J (2011) Can teams help to close the gender competition gap? *Econom. J.* 121(555):1192–1204.

Holt CA, Laury SK (2002) Risk aversion and incentive effects. *Amer. Econom. Rev.* 92(5):1644–1655.

Ifcher J, Zarghamee H (2016) Pricing competition: A new laboratory measure of gender differences in the willingness to compete. *Experiment. Econom.* 19(3):642–662.

Knight J (2002) Sexual stereotypes. *Nature* 415(6869):254–256.

Lichtenstein S, Fischhoff B, Phillips LD (1982) Calibration of probabilities: The state of the art to 1980. Daniel K, Slovic P, Tversky A, eds. *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge University Press, Cambridge, UK), 306–334.

Möbius MM, Niederle M, Niehaus P, Rosenblat TS (2011) Managing selfcConfidence: Theory and experimental evidence. NBER Working Paper w17014, National Bureau of Economic Research, Cambridge, MA.

Niederle M, Vesterlund L (2007) Do women shy away from competition? Do men compete too much? *Quart. J. Econom.* 122(3): 1067–1101.

Niederle M, Vesterlund L (2010) Explaining the gender gap in math test scores: The role of competition. *J. Econom. Perspect.* 24(2): 129–144.

Niederle M, Vesterlund L (2011) Gender and competition. *Annual Rev. Econom.* 3:601–630.

Niederle M, Segal C, Vesterlund L (2012) How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Sci.* 59(1):1–16.

Petrie R, Segal C (2015) Gender differences in competitiveness: The role of prizes. GMU Working Paper in Economics 14-47, George Mason University, Fairfax, VA.

Rigdon ML (2012) An experimental investigation of gender differences in wage negotiations. Working paper, Rutgers University, New Brunswick, NJ.

Shurchkov O (2012) Under pressure: Gender differences in output quality and quantity under competition and time constraints. *J. Eur. Econom. Assoc.* 10(5):1189–1213.

Soll JB, Klayman J (2004) Overconfidence in interval estimates. *J. Experiment. Psych. Learn. Memory Cognition* 30(2):299–314.

Sutter M, Ruützler D (2014) Gender differences in competition emerge early in life and persist. *Management Sci.* 61(10): 2339–2354.

Trivers RL (1972) Parental investment and sexual selection. Campbell B, ed. *Sexual Selection and the Descent of Man* (Aldine, Chicago), 136–177.

Vandegrift D, Yavas A (2009) Men, women, and competition: An experimental test of behavior. *J. Econom. Behav. Organ.* 72(1): 554–570.

Wozniak D, Harbaugh WT, Mayr U (2014) The menstrual cycle and performance feedback alter gender differences in competitive choices. *J. Labor Econom.* 32(1):161–198.