Do women give up competing more easily? Evidence from the lab and the Dutch Math Olympiad

Buser, T.; Yuan, H.

DOI
10.2139/ssrn.2867346
10.1257/app.20170160

Publication date
2019

Document Version
Final published version

Published in
American Economic Journal: Applied Economics

American Economic Journal: Applied Economics 2019, 11(3): 225–252
https://doi.org/10.1257/app.20170160

Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad†

By Thomas Buser and Huaiping Yuan*

We use lab experiments and field data from the Dutch Math Olympiad to show that women are more likely than men to stop competing if they lose. In a math competition in the lab, women are much less likely than men to choose competition again after losing in the first round. In the Math Olympiad, girls, but not boys, who fail to make the second round are less likely to compete again one year later. This gender difference in the reaction to competition outcomes may help to explain why fewer women make it to the top in business and academia. (*JEL* C90, D82, D91, J16)

People who want to advance in their careers are regularly confronted with situations that are in essence winner-takes-all competitions, including admission procedures to high-ranked universities, job applications, promotion tournaments, and competition between entrepreneurs. Whether an individual actively seeks out such situations or tries to avoid them will partially determine what kind of career they end up in and how far they advance up the hierarchy. Starting with Gneezy, Niederle, and Rustichini (2003) and Niederle and Vesterlund (2007), a large literature in experimental economics shows that women are often less willing to compete than men. This gender gap has attracted attention as a potential explanation for gender differences in career choices and labor market outcomes (Bertrand 2011).

In most experimental studies, the decision to compete or not is one-shot. But career decisions are typically made in a dynamic setting. To make it to the top, one has to not only be willing to compete but be willing to keep competing after a setback. Observational data indicates that in male-dominated fields, such as science, engineering, or economics, women are not only less likely to enter but also more likely to drop out (European Commission 2013; Ellis, Fosdick, and Rasmussen 2016; Price 2010). In particular, women in these fields have been found to be more likely than men to drop out after experiencing setbacks or competitive pressure early

* Buser: School of Economics, University of Amsterdam and Tinbergen Institute, Roetersstraat 11, 1018WB Amsterdam, The Netherlands (email: t.buser@uva.nl); Yuan: School of Economics, University of Amsterdam, Roetersstraat 11, 1018WB Amsterdam, The Netherlands (email: h.yuan@uva.nl). Alexandre Mas was coeditor for this article.We would like to thank seminar and workshop participants in Amsterdam, Bilbao, Cologne, Hong Kong, Leuven, Maastricht, Munich, San Francisco, São Paulo, Stanford, Tel Aviv, Tilburg, Vienna, and Zürich for useful comments. Thomas Buser gratefully acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO) through a personal Veni grant (451-14-022) and we would like to thank CREED for letting us use their lab.

† Go to https://doi.org/10.1257/app.20170160 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

on in their careers (Goldin 2013, Felder et al. 1995, Fischer 2017, Goodman et al. 2002, Katz et al. 2006).

We conduct three incentivised lab experiments and use field data from the Dutch Math Olympiad to empirically investigate gender differences in willingness to compete in a dynamic setting. In particular, we ask whether there is a gender difference in the propensity to stop competing after losing.

In our lab experiments, participants perform the same task over a number of rounds. In each round, they decide whether they want to receive a piece-rate payment or enter a winner-takes-all competition against a randomly selected opponent. Participants who enter the competition and outperform their opponent receive twice the piece rate, those who enter the competition and lose receive no payment in that round. After each round, participants who choose to compete learn whether they won or lost against their opponent. By comparing the likelihood of choosing competition in subsequent rounds between participants who win in round 1 and participants who lose in round 1, we can determine the effect of losing and whether there is a gender difference in the reaction to losing. To identify the causal effect of losing versus winning, we use the fact that conditional on a participant's own performance, the competition outcome is exogenous. Our main result is that women are much more likely than men to stop competing if they lose in the first round.

The Dutch Math Olympiad is a yearly math competition that draws participants mainly from years four and five of secondary school. Our sample of fourth-year students who participated in 2010–2014 contains more than 11,000 observations. In each year, a fixed number of top performers advance to the second round of the competition. All those who participate in year four can participate again in year five regardless of their performance. This generates a sharp regression discontinuity design where we can compare participants who score just above and just below the cutoff score needed to advance to the second round to determine the causal effect of losing on the likelihood of participating again a year later, and the gender difference therein. We find that girls who do not make the cutoff are substantially less likely to participate again a year later while there is no effect for boys.

The seminal study on gender differences in the willingness to enter competitions is Niederle and Vesterlund (2007). In a lab experiment, they find that, conditional on ability, women are much less likely than men to enter a winner-takes-all math competition. Strikingly, this gender difference is present also among high performers, who hurt their expected earnings by not competing.[1] A growing literature confirms the external relevance of this result, either by demonstrating that competition decisions made in the lab predict career choices (Buser, Niederle, and Oosterbeek 2014; Zhang 2012; Berge et al. 2015; Reuben, Sapienza, and Zingales 2015; Reuben, Wiswall, and Zafar 2017; Buser, Geijtenbeek, and Plug 2018; Buser, Peter, and

[1] This result has been replicated numerous times (see Croson and Gneezy 2009, Niederle and Vesterlund 2011, and Niederle 2016 for surveys). A separate literature started by Gneezy, Niederle, and Rustichini (2003) shows that men's performance tends to increase more strongly in response to competitive pressure than women's, but a large number of studies on tournament entry indirectly show that this is not the case for the task used in the present experimental design (Niederle and Vesterlund 2011). Iriberri and Rey-Biel (forthcoming) shows that this effect is present in the context of a mathematics contest for students aged 10 to 16: girls underperform relative to their math grades and this underperformance is worse in later stages of the contest, where competitive pressure is higher.

Wolter 2017a, b) or by recruiting people for real jobs and showing that compensation schemes that depend on relative performance deter women from applying relative to men (Flory, Leibbrandt, and List 2015; Samek 2015).

How the gender gap in willingness to compete evolves in response to winning and losing remains an open question. Starting with Niederle and Vesterlund (2007), many experiments show that women are less confident than men. If the gap is mainly due to wrong beliefs, and if winning and losing affect choices exclusively through a feedback effect, gender differences should shrink in response to experience as long as people rationally update their beliefs. A number of papers have studied the effect of feedback on the gender gap in willingness to compete. Ertac and Szentes (2011) and Wozniak, Harbaugh, and Mayr (2014) show that providing precise relative performance feedback prior to choosing the payment scheme can eliminate the gender gap in willingness to compete, while Cason, Masters, and Sheremeta (2010) finds a significant gender gap even after providing relative feedback.

A number of experimental studies suggest, however, that winning or losing in competitions has effects above mere informational value, and that men and women may react differently to winning and losing in ways that may enhance the gender gap. Buser (2016) finds that male competition losers become more challenge seeking whereas female losers subsequently perform worse at the same task. Gill and Prowse (2014) similarly finds that women are more prone than men to reduce their performance after losing a competition. Cai et al. (forthcoming) finds a similar effect in a highly competitive university entrance exam in China, where women, but not men, perform worse on afternoon exams if they performed worse than expected on another exam in the morning of the same day.[2] Möbius et al. (2011); Buser, Gerhards, and van der Weele (2018); and Coutts (forthcoming) all find that men update their beliefs about their own abilities more strongly in response to both positive and negative feedback, which could perpetuate initial gender differences in beliefs for high performers, and Berlin and Dargnies (2016) finds gender differences in how feedback is processed in a competitive setting. Sutter and Glätzle-Rützler (2015) twice measure willingness to compete in the same sample of teenagers with two years in between measurements and find that girls are more likely than boys to switch from competition to piece rate, while boys are more likely to switch from piece rate to competition.

None of these studies directly address the question of whether there is a gender difference in the effect of winning or losing in a competition on subsequent willingness to compete again. In career settings, actually competing—and consequently winning

[2] A number of papers use data from professional sports to analyze gender differences in the effect of competition outcomes. These are subject to the caveats that mixed-gender competition is rare, samples are highly selective, and identification of gender differences is complicated by the fact that the competitive settings and incentives are rarely equal for male and female athletes even within the same sport. The most relevant is Wozniak (2012), who finds that the effect of prior tournament outcomes on the propensity of entering further tournaments is similar for male and female tennis players (he does find a subtle gender difference, though, in how long the effect of prior experience lasts, with men being more affected than women by outcomes further in the past). Also looking at tennis players, De Paola and Scoppa (2017) finds that female, but not male, players perform worse after losing the first set. Legge and Schmid (2013) finds that skiers who narrowly miss a place on the podium underperform in the next race, this effect being somewhat stronger for female skiers, and Rosenqvist (2016) finds that golfers who narrowly miss the cut underperform in the next tournament but does not find a gender difference.

or losing—is often the only way to receive information on relative performance. Opportunities to receive detailed feedback prior to the decision of entering a certain competition are rare. Moreover, only people who are willing to enter a competition in the first place learn something about their own relative ability. Our experiments are designed to reflect these features.

Our combined results from the lab and the field show that winning and losing have effects on willingness to compete above mere provision of relative performance feedback and provide robust evidence for a gender difference in the reaction to losing. The rest of the paper is structured as follows. Section I explains the experimental design. Section II presents the experimental results and explores potential mechanisms. Section III describes the results from the field study. Section IV concludes by discussing potential implications for our understanding of gender differences in career choices and labor market outcomes.

I. Experimental Design

Our experimental design is based on Niederle and Vesterlund (2007). Participants earn money for their performance in a real-effort task that consists of adding up sets of five two-digit numbers. We ran three experiments, which we will refer to as the "main," "feedback," and "risk" experiments. We will start by describing the design of the main experiment and then lay out the differences in the other two experiments.

The main experiment consists of six paid rounds. Participants have three minutes per round to solve as many addition problems as they can. Wrong answers are not penalized. In each round, participants are paired with a new randomly chosen, anonymous opponent and have to choose how they would like to be paid for their performance. They can choose between piece-rate and competitive remuneration. If they choose the piece rate, they receive one point per correct answer regardless of the performance of their opponent. If they choose competition, they receive two points per correct answer if they score higher than their opponent and nothing if they do not (in case of a tie, winning or losing is randomly determined). Opponents are randomly chosen amongst all other participants regardless of their choice. One point is worth 25 euro cents and all rounds are paid.

After each round, participants who choose competition receive feedback on their absolute and relative performance. That is, they learn their score and whether they won or lost against their opponent (the feedback reads "You scored X correct answers. You scored higher (lower) than your opponent. You therefore won (lost) against your opponent."). Before the start of each round, we elicit an incentivised measure of participants' beliefs about their own performance. Participants are asked to predict their rank compared to all other participants in their session in the upcoming round. If their guess is within plus-minus one of the true rank, they receive a bonus of four points.

Before the start of the incentivised part of the experiment, there is a three-minute practice round after which the participants learn their score but receive no relative feedback. At the end of the last round, before they see a screen that summarizes their earnings, participants fill in a short questionnaire asking for their gender and age, and their perception about their own willingness to take risk and their own competitiveness.

For an unincentivised measure of risk attitudes, we ask "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" (Dohmen et al. 2011). The answer is on a scale from 0 ("unwilling to take risks") to 10 ("fully prepared to take risk"). For an unincentivised measure of competitive attitudes, we ask an analogous question, "How competitive do you consider yourself to be? Please choose a value on the scale below, where the value 0 means 'not competitive at all' and the value 10 means 'very competitive.'"

To summarize, in each of the six rounds the timeline is as follows. Participants:

- predict their rank in the upcoming round compared to all other participants present in the lab;
- choose between the competitive and piece-rate payment schemes;
- perform in the task;
- learn their score and, if they chose competition, whether or not they beat their opponent.

In the feedback experiment, we ask what happens if those who choose the piece rate receive feedback that is equivalent to the feedback received by those who compete. In this experiment, participants learn whether they outperformed their opponent even if they chose the piece rate (for participants who chose the piece rate, the message reads "You scored X correct answers. You scored higher (did not score higher) than your opponent. You therefore would have won (lost) against your opponent."). The rest of the design is identical to the main experiment except that there are four rounds of four minutes each, whereas the main experiment consists of six rounds of three minutes each.[3]

The aim of the risk experiment is to determine whether we find the same patterns in a setting that is identical to the main experiment but in which winning and losing are entirely due to luck. Participants perform the adding-up-numbers task over six rounds of three minutes. In each round, they decide between a safe piece rate of 1 point per correct answer and random remuneration, whereby they receive 2 points per correct answer with probability x and nothing with probability $1 - x$ (x is constant across rounds and randomly fixed at 0.3, 0.4, 0.5, 0.6, or 0.7 with equal probability and participants are informed about their x during the instructions).

In the main experiment, a total of 188 individuals participated in 7 sessions with 21 to 31 participants each. The sessions were run in November 2015. Participants earned 21.2 euros, on average, including a 7 euro show-up fee. In the feedback experiment, a total of 184 individuals participated in 7 sessions with 21 to 31 participants each. The sessions were run in January 2015. Participants

[3] The sessions of the feedback experiment were the first to be run. The reason to use a shorter task duration (three instead of four minutes) for the main experiment was to create greater overlap in scores between first-round winners and losers, which helps when estimating regressions controlling for score dummies, and to be able to observe the effect of first-round outcomes over a larger number of rounds.

Table 1—Descriptive Statistics

		Main experiment				Feedback experiment			
	Scale	All (1)	Male (2)	Female (3)	*p*-value (4)	All (5)	Male (6)	Female (7)	*p*-value (8)
Score per minute	0–6.5	2.43	2.51	2.36	0.243	2.57	2.66	2.47	0.141
		(1.00)	(1.07)	(0.92)		(0.98)	(1.06)	(0.87)	
Belief	0–1	0.59	0.63	0.55	0.012	0.59	0.64	0.53	0.001
		(0.25)	(0.24)	(0.26)		(0.26)	(0.25)	(0.26)	
Choosing competition	Binary	0.47	0.53	0.41	0.046	0.47	0.53	0.40	0.026
		(0.50)	(0.50)	(0.49)		(0.50)	(0.50)	(0.49)	
Earnings per minute	0–3.25	0.76	0.82	0.69	0.102	0.81	0.90	0.72	0.029
		(0.66)	(0.72)	(0.59)		(0.71)	(0.77)	(0.63)	
Risk taking	0–10	5.5	6.0	5.0	0.006	5.4	5.6	5.1	0.282
		(2.45)	(2.33)	(2.47)		(2.76)	(2.68)	(2.84)	
Competitiveness	0–10	6.7	7.1	6.3	0.034	6.8	7.2	6.4	0.035
		(2.39)	(2.20)	(2.51)		(2.56)	(2.52)	(2.54)	
Observations		188	93	95		184	97	87	

Notes: The table shows per-round averages over all rounds. Belief means the probability of winning as predicted by the participants themselves at the start of each round. Earnings are per-minute earnings in euros. Risk-taking and competitiveness are self-rated questionnaire measures. Columns 4 and 8 show *p*-values from *t*-tests of the gender difference.

earned 20.8 euros, on average, including a 7 euro show-up fee. In the risk experiment, a total of 188 individuals participated in 7 sessions with 22 to 31 participants each. Participants earned 18.4 euros, on average, including a 7 euro show-up fee. The sessions were run in June 2016. For all sessions of all experiments, we always opened an equal amount of slots for male and female participants and all sessions were therefore roughly gender-balanced. Students could only participate in one of the experiments. All sessions were run at the CREED lab at the University of Amsterdam and made use of z-Tree (Fischbacher 2007).

II. Experimental Results

In this section, we will describe and analyze our experimental data. We will first describe how the willingness to compete of men and women evolves over the rounds. We will then determine whether there is a gender difference in the effect of losing in the first round on subsequent willingness to compete again. Finally, we will use the data from the risk experiment, as well as the belief elicitation data from the main and feedback experiments, to explore potential mechanisms behind our findings.

A. *Descriptive Results*

Table 1 shows descriptive statistics of average choices in the main and feedback experiments over all rounds by gender. In both experiments, men scored slightly higher compared to women, although this difference is not statistically significant. Men believe they perform significantly better and they are more likely to choose

TABLE 2—GENDER GAP IN COMPETITION ENTRY IN EACH ROUND

	Round 1 (1)	Round 2 (2)	Round 3 (3)	Round 4 (4)	Round 5 (5)	Round 6 (6)	Pooled (2–6) (7)
Panel A. Main experiment							
All							
Female	−0.135	−0.146	−0.091	−0.125	−0.068	−0.085	−0.103
	(0.071)	(0.066)	(0.067)	(0.068)	(0.063)	(0.064)	(0.051)
Observations	188	188	188	188	188	188	940
Comp. in round 1							
Female		−0.214	−0.117	−0.215	−0.043	−0.152	−0.148
		(0.065)	(0.081)	(0.095)	(0.086)	(0.085)	(0.059)
Observations		92	92	92	92	92	460
PR in round 1							
Female		−0.023	0.014	0.041	−0.021	0.030	0.008
		(0.100)	(0.093)	(0.081)	(0.079)	(0.083)	(0.058)
Observations		96	96	96	96	96	480
R1 score	√	√	√	√	√	√	√
R1 rank	√	√	√	√	√	√	√
R1 outcome		(√)	(√)	(√)	(√)	(√)	(√)

	Round 1	Round 2	Round 3	Round 4		Pooled (2–4)
Panel B. Feedback experiment						
All						
Gender	0.012	−0.101	−0.110	−0.115	−0.124	−0.109
	(0.069)	(0.049)	(0.059)	(0.060)	(0.049)	(0.048)
Observations	184	184	184	184	184	552
Comp. in round 1						
Female		−0.255	−0.113	−0.185	−0.173	−0.184
		(0.068)	(0.090)	(0.083)	(0.070)	(0.059)
Observations		87	87	87	87	261
PR in round 1						
Female		0.010	−0.119	−0.101	−0.129	−0.086
		(0.065)	(0.083)	(0.085)	(0.071)	(0.061)
Observations		97	97	97	97	291
R1 score	√	√	√	√	√	√
R1 rank	√	√	√	√	√	√
R1 outcome		√	√	√	√	√
Outcomes R2–3					√	

Notes: The table shows coefficients from OLS regressions of a binary indicator for choosing the competition on a gender dummy. Each line is a separate regression: R1 score, R1 rank, and R1 outcome means score, normalized within-session rank and the competition outcome in round 1. In panel A, only the regressions for those who compete in round 1, and therefore receive feedback, control for the round 1 outcome. All regressions control for session fixed effects. Robust standard errors are in parentheses. Standard errors for the pooled regressions are clustered at the individual level. The main experiment consisted of six rounds of three minutes each and the feedback experiment consisted of four rounds of four minutes each.

competition over piece rate both in the main and in the feedback experiment.[4] The earnings of men are slightly higher than the earnings of women in both experiments. In the questionnaires, men judge themselves to be more risk-taking and more competitive.

In Table 2, we show the gender gap in the likelihood of choosing competition in each round of the main and feedback experiments. We use OLS regressions of a dummy for choosing competition in a given round on gender, controlling for absolute

[4] Beliefs are expressed as the probability of winning as predicted by the participants themselves and are based on the rank predictions of the participants. The variable is calculated as (predicted number of participants with lower score)/(number of participants − 1).

performance and chance of winning[5] (the raw tournament entry rates of men and women over the rounds are visualized in Figures A1 to A3 in the online Appendix). We use absolute performance and chance of winning in round 1 as controls for all rounds. Because the choice of remuneration scheme might itself affect performance levels (for example, the same participant might try harder in those rounds where she chooses to compete), and because performance may be affected by earlier competition outcomes, controlling for later-round scores could bias the estimate of the gender coefficient.[6] We explore the importance of changes in relative performance over the rounds (and gender differences therein) as a mechanism behind our results in Section IIC.

Columns 1 to 6 of panel A in Table 2 show the gender difference in each of rounds 1 to 6 of the main experiment. Men are more likely than women to choose competition over piece rate in all rounds. The gender difference is 14 percentage points initially and fluctuates between 7 and 15 percentage points over rounds 2 to 6. Panel A also shows separate regressions for the gender gap in rounds 2 to 6 for those who choose competition in round 1 and those who choose piece rate in round 1. The most striking result is that even for those people who compete in round 1, a large gender difference appears in round 2 of the main experiment. In round 2, a woman who chooses competition in round 1 is 21 percentage points less likely to compete again than a man with the same round 1 choice, same absolute performance, same chance of winning, and same competition outcome in round 1. The gender difference is 12 and 22 percentage points in rounds 3 and 4, and is still 15 percentage points in the last round. The pooled regression in column 7 shows that the average woman who chooses to compete in round 1 is 15 percentage points less likely to compete in each of the following 5 rounds compared to the average man who competes in round 1. For those who initially choose the piece rate, the gender difference is close to zero in all subsequent rounds.

Panel B repeats this analysis using data from the feedback experiment, where non-competers learn whether they would have won or lost had they chosen competition. Columns 1 to 4 show regression results for rounds 1 to 4. Column 5 repeats the analysis for round 4 while additionally controlling for the competition outcomes in rounds 2 and 3. The coefficient therefore shows the difference in tournament entry in round 4 between men and women who have the same initial performance and rank and who received the same feedback in rounds 1 to 3.[7] The gender difference in the likelihood of choosing to compete is initially close to zero. But this grows to around 10 percentage points in round 2 and 12 percentage points in round 4.[8]

[5] Chance of winning is calculated as (number of participants with lower score)/(number of participants − 1) and therefore corresponds to rank normalized by the number of subjects in a session.

[6] When using round 1 performance as a control, the problem of the remuneration scheme possibly affecting performance still applies in the case of the regressions that use the whole sample (which contains both people who compete in round 1 and people who choose the piece rate in round 1). However, the problem is completely avoided in case of the regressions using the subsamples of those who compete in round 1 and those who choose the piece rate in round 1.

[7] Controlling for outcomes in rounds 2 and higher is only possible in the feedback experiment where even those participants who stop competing learn whether they "won" or "lost" after each performance.

[8] A potential reason for the small initial gender gap is that, contrary to the main experiment (and contrary to the standard design in the literature on gender differences in willingness to compete), it is no longer possible to avoid relative feedback by staying away from competition. Women have been found to be more feedback-averse

For those who initially choose competition, the pattern in the feedback experiment looks remarkably similar to the main experiment. The pooled regression in column 6 shows that the average woman who competes in round 1 is 18 percentage points less likely to compete over each of rounds 2 to 4 than the average man who competes in round 1. For those who choose the piece rate in round 1, a gender gap of around 11 percentage points emerges in rounds 3 and 4, although this is only statistically significant at the10 percent level in round 4 when additionally controlling for the feedback received in rounds 2 and 3 (column 5).[9]

B. *The Gender Difference in the Effect of Losing on Subsequent Willingness to Compete*

We will now investigate whether the larger proportion of women who drop out of competition and the resulting emergence of a gender gap for those who initially compete are due to a gender difference in the reaction to competition outcomes. In particular, we will answer the question whether women who initially choose to compete are more likely than their male counterparts to stop competing following a loss.

It is important to note that conditional on performance, winning or losing in round 1 only depends on the score of the randomly allocated opponent and can consequently be interpreted as a random treatment. The regressions in this section control for score fixed effects and therefore give us causal estimates of the effect of losing relative to winning on the likelihood of competing in the subsequent rounds. The analysis will focus on the reaction to the competition outcome in round 1. The reason for this focus is that in later rounds, the participants who choose to compete have heterogeneous prior experience and are selected on their reaction to earlier competition outcomes. The effect of competition outcomes in later rounds can therefore not be interpreted in a causal way.

Figure 1 shows how often, on average, participants choose to compete from round 2 onward, split by their gender, their choice of remuneration scheme in round 1, and whether they won or lost in round 1.[10] Panel A shows competition rates over rounds 2 to 6 of the main experiment for men and women who choose to compete in round 1. Participants of both genders who win the competition in round 1 mostly keep competing over the following five rounds (4.6 out of 5 times for men and 4.2 out of 5 times for women; $p = 0.17$, t-test). However, the reaction to losing in

than men, on average (Möbius et al. 2011), which might encourage choosing the piece rate if this allows to avoid relative feedback.

[9] Another noteworthy difference is that for those who choose piece rate in round 1, a smaller proportion than in the main experiment decide to enter in round 2, and that competition entry for this group keeps increasing in rounds 3 and 4, whereas in the main experiment, it steadily declines after its peak in round 2 (see Figure A3 in the online Appendix). A likely explanation for this difference is that, lacking feedback, more low-performing people choose to enter competition after round 1 in the main experiment, whereas in the feedback experiment, only those who receive positive feedback choose to enter. This is confirmed by the data. The mean chance of winning (calculated based on round 1 scores) of those who choose to enter in round 2 is 61 percent in the feedback experiment compared to 48 percent in the main experiment.

[10] Fifty-two men and 40 women chose to compete in round 1 of the main experiment, of whom 18 and 14 lost, respectively. Forty-eight men and 39 women chose to compete in round 1 of the feedback experiment, of whom 20 and 18 lost, respectively. Forty-nine men and 48 women chose the piece rate in round 1 of the feedback experiment, of whom 17 and 14 would have won, respectively.

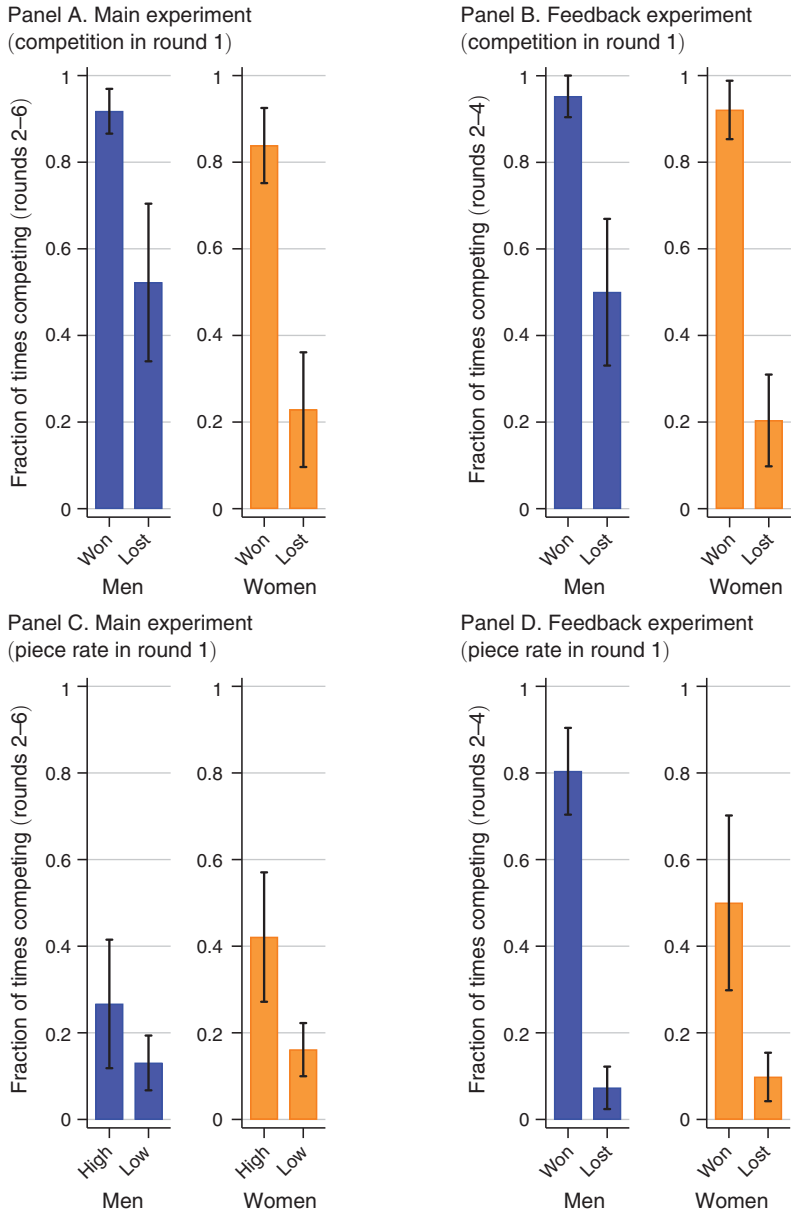

Figure 1. Average Number of Times Competition Is Chosen after Round 1 by Gender, Initial Choice, and Competition Outcome in Round 1

Notes: The bars show the fraction of rounds that participants chose to compete over rounds 2 to 6 (2 to 4 in case of the feedback experiment). "Lost" denotes participants who lost in round 1 and "Won" denotes participants who won in round 1. "High" denotes participants who performed in the top 50 percent in round 1 and "Low" denotes participants who scored in the bottom 50 percent in round 1. Error bars represent 90 percent confidence intervals. The main experiment consisted of six rounds of three minutes each and the feedback experiment consisted of four rounds of four minutes each.

round 1 differs strongly between men and women. Men who initially compete and lose compete a further 2.6 out of 5 times, whereas for women this is only 1.1 times ($p = 0.04$).

Panel B repeats this analysis for the feedback experiment. The pattern is remarkably similar to the main experiment. Both men and women essentially keep competing if they win in round 1 (2.9 out of 3 times for men and 2.8 out of 3 times for women; $p = 0.50$, t-test). But women who lose are again much less likely to compete over the following rounds than men who lose (1.5 out of 3 times for men and 0.6 out of 3 times for women; $p = 0.02$).

Giving feedback to non-competers in the feedback experiment allows us to also test whether there is a gender difference in the reaction to positive feedback for those who are initially not willing to compete. Panel D of Figure 1 shows the average number of times participants choose competition over rounds 2 to 4 for those who choose piece rate in round 1, separately by gender and the feedback received at the end of round 1. There is no gender difference for those who perform worse than their opponent in round 1: both men and women essentially keep choosing the piece rate (men compete 0.2 out of 3 times and women 0.3 out of 3 times; $p = 0.57$). But men who outperform their opponent compete more often over the following rounds than women who outperform their opponent (2.4 out of 3 times versus 1.5 out of 3 times; $p = 0.02$). We cannot repeat this analysis for the main experiment as those who choose piece rate in the main experiment do not receive feedback. As a comparison point, in panel C, we split the sample from the main experiment into those with a higher and lower than 50 percent chance of winning.

In Table 3, we use OLS regressions that estimate the gender difference in the reaction to losing conditional on absolute and relative performance (score fixed effects and likelihood of winning in round 1). We regress a binary indicator for choosing to compete on a gender dummy, a binary indicator for having lost in round 1, and the interaction of the two. Each individual decision in rounds 2 to 6 is a separate observation and standard errors are clustered at the individual level.

The regression in column 1 of Table 3 uses the sample of participants in the main experiment who choose to compete in round 1. The results show that the gender difference in the reaction to losing is robust to controlling for absolute and relative performance. Whereas male losers are 24 percentage points less likely to enter competition in rounds 2 to 6 compared to winners, this effect is more than twice as large for women at 59 percentage points. In column 3, we repeat this analysis for the feedback experiment. The coefficients look very similar. While male losers are 28 percentage points less likely to choose competition in rounds 2 to 4 compared to male winners, this effect is 56 percentage points for women.[11]

It is interesting to check whether the gender difference in the reaction to losing applies also to high-performing individuals who have a positive expected return from competing. If this is the case, losing early on is more costly for high-performing women than for high-performing men. In columns 2 and 4, we restrict the sample

[11] In Table A1 in the online Appendix, we pool the data from the main and feedback experiments and interact all variables with experiment dummies. This analysis confirms that the gender difference in the effect of losing on subsequent competition choices is very similar across the two experiments.

Table 3—Effect of Competition Outcome in Round 1 on Subsequent Competition Entry

	Competition in round 1					Piece rate in round 1	
	Main experiment		Feedback experiment			Feedback experiment	
	All	Top	All	Top		All	Top
	(1)	(2)	(3)	(4)		(5)	(6)
Female	−0.065	−0.132	−0.028	−0.028	Female	0.028	0.213
	(0.056)	(0.053)	(0.073)	(0.073)		(0.049)	(0.161)
Round 1 lose	−0.240	−0.149	−0.284	−0.171	Round 1 win	0.564	0.645
	(0.126)	(0.156)	(0.124)	(0.145)		(0.092)	(0.109)
Female × lose	−0.349	−0.409	−0.273	−0.332	Female × win	−0.287	−0.415
	(0.144)	(0.186)	(0.129)	(0.194)		(0.115)	(0.196)
Score FE	√	√	√	√		√	√
Round 1 rank	√	√	√	√		√	√
Observations	460	275	261	177		291	96
Individuals	92	55	87	59		97	32

Notes: The table shows coefficients from OLS regressions of a binary indicator for choosing the competition in rounds 2 to 6 (2 to 4 in case of the feedback experiment) on a gender dummy, a dummy for having lost the round 1 competition, and the interaction of the two. Each choice is a separate observation and standard errors are clustered at the individual level. The sample in columns 1 to 4 consists of participants who choose competition in round 1. The sample in columns 5 and 6 consists of participants who choose piece rate in round 1. The columns marked "Top" restrict the sample to participants who have a higher than 50 percent chance of winning based on their round 1 performance. Score fixed effects and round 1 rank mean score and normalized within-session rank in round 1. Clustered standard errors are in parentheses. The main experiment consisted of six rounds of three minutes each and the feedback experiment consisted of four rounds of four minutes each.

to those participants who have a higher than 50 percent chance of winning.[12] In the main experiment, high-performing men hardly react to losing in round 1 but high-performing women who lose are 56 percentage points less likely to choose competition again over the subsequent rounds compared to high-performing women who win. This result is mirrored in the feedback experiment.

We will now look at the sample of participants who choose the piece rate in the first round of the feedback experiment. The regression results in columns 5 and 6 of Table 3 show that the gender difference in the effect of a positive outcome on future willingness to compete, for those who are initially not willing to compete, is significant conditional on performance. Whereas men who choose the piece rate and "win" are 56 percentage points more likely to compete over the following three rounds than men who "lose;" this effect is only 28 percentage points for women. This gender difference is even bigger for top performers with a higher than 50 percent chance of winning.[13]

[12] Chance of winning is calculated within session based on first-round performance as (number of participants with lower score)/(number of participants − 1).

[13] When, analogous to Figure 1, we run a regression for the sample of those who choose the piece rate in round 1 of the main experiment, where we use an indicator for being in the top 50 percent of performers instead of a winner/loser dummy, the interaction effect goes in the opposite direction, confirming the impact of feedback (the coefficient on the interaction term is equal to 0.167 with a standard error of 0.142). In Table A1 in the online Appendix, we pool the data from the main and feedback experiments and interact all variables with experiment dummies. This analysis confirms that the gender difference in the effect of winning/being top on subsequent competition choices is significantly different across the two experiments.

Our analysis so far documents the effect of losing in round 1 on average choices over all subsequent rounds. It is also interesting to ask how this effect changes over the rounds, taking full advantage of the experimental data. Figure A4 in the online Appendix shows the proportion of men and women who choose to compete in each round by round 1 outcome (winning or losing) and round 1 choice (competition or piece rate) for the pooled sample from the main and feedback experiments. The main insight is that the gender difference in the reaction to losing for those who compete in round 1 is strongest in round 2, and then attenuates over the rounds but never completely disappears. The figure also reveals an interesting pattern for participants who choose the piece-rate in round 1 of the feedback experiment. Rather than men being more likely to start competing after "winning," the gender difference in the reaction to winning amongst this group is due to women being more likely to drop out of competition again in round 3.

In expectation, competition entry is profitable for high-performing individuals, while it is costly for low-performing individuals. Our results show that the gender difference in the reaction to losing occurs both for low and high performers. In Figure A5 in the online Appendix, we quantify the lost earnings from wrong decisions for men and women at different parts of the performance distribution. In particular, we show the difference between actual expected earnings given performance and competition choice and the expected earnings from the optimal choice given performance. The analysis pools the data from rounds 2 to 4 of the main and feedback experiments and splits the sample into performance quartiles. Among participants who perform in the upper two quartiles in a given round, women, by competing less, in expectation forgo a higher amount than men, whereas for participants in the bottom half, men forgo slightly more money than women. The difference is largest for the top 25 percent performers (57 euro cents lost per round for women versus 35 euro cents for men). At the bottom end of the performance scale, the gender difference is reversed (25 euro cents for men and 15 euro cents for women), reflecting that low-performing men enter competition at higher rates.

Our experiments were not designed to estimate the effect of losing versus winning in later rounds. Even conditional on the choice in round 1, competing in later rounds is endogenous as it is a result of individual reactions to previous competition outcomes. In the online Appendix, we present two analyses where we nevertheless try to learn as much from the data as we can. Both approaches, however, are based on fairly small subsamples. The first approach is to restrict the sample in each round to participants who chose to compete and won in all previous rounds (Figure A6 and Table A2). We do not observe a gender difference in the reaction to losing in round 2, but we observe a sizable and significant gender difference in round 3. The second approach takes advantage of the fact that in the feedback experiment, participants learn about the competition outcome in each round regardless of their choice. Figure A7 shows the average choices of men and women by their choice in round 1 (competition or piece rate) and the competition outcomes in rounds 1 and 2.[14] The

[14] This creates groups of participants that are homogeneous in their received feedback even if they may take different decisions in round 2. Note that we cannot repeat this exercise for the main experiment because, there, learning about the competition outcome in any round is conditional on choosing competition.

most interesting observation comes from those participants who choose the piece rate in round 1: amongst those who win in round 1 and then lose in round 2, 83 percent of men, but only 17 percent of women, compete in rounds 3 and 4. This shows that the gender difference in the reaction to positive feedback observed in Table 3 and Figure 1 is really a gender difference in the reaction to losing in round 2 after winning in round 1. While imperfect, these two analyses combined suggest that competition outcomes in later rounds matter too.

C. *The Role of Risk Preferences*, *Initial Beliefs*, *Belief Updating*, *and Performance*

The experimental economics literature on gender differences in willingness to compete suggests a number of potential mechanisms behind the gender difference in the reaction to competition outcomes, namely gender differences in risk preferences, initial beliefs, belief updating, and the evolution of relative performance over time. Our experimental data allow us to explore all four of these.

Risk.—Women are often found to be more risk averse than men (Croson and Gneezy 2009, Charness and Gneezy 2012) and recent papers by Gillen, Snowberg, and Yariv (2015) and van Veldhuizen (2016) indicate that gender differences in risk preferences might play an important role in explaining the gender gap in tournament entry in a static setting. In particular, van Veldhuizen (2016) lets participants choose between piece-rate remuneration and a risky option, where winning and losing is determined by a lottery. He finds a large and significant gender gap in choosing the risky option.

In this section, we will use the data from our risk experiment to explore whether changes in risk preferences can explain the dynamic patterns in willingness to compete observed in the main and feedback experiments. The setting of the risk experiment is identical to the main experiment but winning and losing are determined by a random draw. This allows us to investigate whether the main result—women are more likely than men to stop competing after losing—occurs also in an environment that is characterized by the same risks, but which lacks the competitive dimension.[15]

Table 4 reproduces the regressions in Table 2, using the data from the risk experiment, showing the gender difference in the likelihood of choosing the risky option in each round conditional on performance and exogenous probability of winning. Whereas men are 21 percentage points more likely to choose the risky option in round 1, this gender gap is smaller and not consistently significant over the remaining rounds. Looking separately at participants who choose the risky and safe options in round 1 reveals differences with the patterns observed in the main and feedback experiments. For those who choose the risky option in round 1, no significant gap appears in round 2. However, a gap starts appearing from round 3 onward and

[15] Over all rounds of the risk experiment, men performed slightly but not significantly better than women and earned slightly but not significantly more. Again, men judged themselves to be more risk-seeking and competitive. However, men were not significantly more likely to choose the risky option compared to women over all six rounds. Scores per minute: 2.47 (men) versus 2.32 points (women); $p = 0.21$, t-test. Earnings per minute: 0.66 versus 0.61; $p = 0.31$. Risk taking: 5.6 versus 4.5; $p = 0.00$. Competitiveness: 7.4 versus 6.8; $p = 0.06$. Choosing the risky option: 0.38 versus 0.34; $p = 0.48$.

Table 4—Gender Gap in Choosing Risky Option in Each Round

	Round 1 (1)	Round 2 (2)	Round 3 (3)	Round 4 (4)	Round 5 (5)	Round 6 (6)	Pooled (2–6) (7)
All							
Female	−0.207	−0.027	−0.116	0.041	−0.107	0.014	−0.039
	(0.058)	(0.064)	(0.063)	(0.064)	(0.061)	(0.055)	(0.041)
Observations	188	188	188	188	188	188	940
Risky in round 1							
Female		−0.049	−0.113	−0.226	−0.180	−0.168	−0.147
		(0.126)	(0.114)	(0.122)	(0.124)	(0.119)	(0.083)
Observations		62	62	62	62	62	310
Safe in round 1							
Female		0.122	0.025	0.277	0.085	0.218	0.145
		(0.071)	(0.071)	(0.063)	(0.070)	(0.050)	(0.034)
Observations		126	126	126	126	126	630
R1 score	√	√	√	√	√	√	√
Pr FE	√	√	√	√	√	√	√
R1 outcome		(√)	(√)	(√)	(√)	(√)	(√)

Notes: The table shows coefficients from OLS regressions of a binary indicator for choosing the risky option on a gender dummy. Each line is a separate regression. R1 score and R1 outcome means score and lottery outcome in round 1. Only the regressions for those who choose the risky option in round 1, and therefore receive feedback, control for the round 1 outcome. All regressions control for session fixed effects. Probability fixed effects (Pr FE) means dummy variables indicating the probability of winning the lottery (randomly fixed at 0.3, 0.4, 0.5, 0.6, or 0.7). Robust standard errors are in parentheses. Standard errors from the pooled regressions are clustered at the individual level.

women who start out choosing the risky option are significantly less likely than men to do so again over the remaining rounds. The pattern for those who choose the safe option in round 1 also looks different from the patterns in the other experiments. From round 2 onward, women in this group are more likely to choose the risky option than are men in this group.

In Table 5, we repeat the analysis in Table 3, and use OLS regressions to estimate the gender difference in the reaction to losing conditional on probability of winning and performance. Here, we regress a binary indicator for choosing the risky option in rounds 2 to 6 on a gender dummy, a binary indicator for having lost in round 1, and the interaction of the two. The regressions therefore give us the gender difference in the reaction to losing, conditional on the exogenous probability of winning and performance. The coefficient on the gender interaction is very close to zero. Women are no more likely than men to stop choosing the risky option if they lose in round 1. This demonstrates that the gender difference in the reaction to competition outcomes observed in the previous two experiments does not occur in an environment that is characterized by the same risk structure but lacks the competitive element (competing and winning or losing against someone else's performance). This indicates that gender differences in the dynamic evolution of risk preferences are unlikely to be a mechanism behind the gender difference in the reaction to losing in a competition.

Beliefs and Performance.—We can use the belief-elicitation and performance data from the main and feedback experiments to check whether gender differences in initial beliefs, belief updating, or the evolution of relative performance are

Table 5—Choosing the Risky Option in Rounds 2 and 6

	All observations (1)	Chance of winning ≥ 0.5 (2)
Female	−0.194	−0.220
	(0.112)	(0.123)
Round 1 loser	−0.098	−0.122
	(0.085)	(0.090)
Female × loser	0.011	0.024
	(0.167)	(0.163)
Round 1 score	√	√
Probability FE	√	√
Observations	310	275
Individuals	62	55

Notes: The table shows coefficients from OLS regressions of a binary indicator for choosing the risky option in rounds 2 to 6 on a gender dummy, a dummy for having lost the round 1 random draw, and the interaction of the two. Each choice is a separate observation and standard errors are clustered at the individual level. The sample consists of participants who choose the risky option in round 1. Probability fixed effects means dummy variables indicating the probability of winning the lottery (randomly fixed at 0.3, 0.4, 0.5, 0.6, or 0.7). Clustered standard errors are in parentheses.

mechanisms behind the gender difference in the reaction to competition outcomes. Initial beliefs are measured by the participants' predicted rank at the start of round 1, belief updating is measured by the change in predicted rank over time, and relative performance is measured by the change in actual rank over time.

These mechanisms are explored in Table 6 using OLS regressions. Column 1 replicates the analyses in columns 1, 3, and 5 of Table 3. We then control for each potential mechanism separately in columns 2 to 4. If a mechanism has explanatory power for the gender difference in the reaction to competition outcomes, the magnitude of the gender-interaction coefficient should shrink.

Most studies on gender differences in competitiveness find that, conditional on performance, women tend to be less confident than men (Niederle and Vesterlund 2011), which is also the case in our sample.[16] A loss could therefore more easily push women's beliefs below the threshold at which they are no longer willing to compete. This is explored in column 2 by controlling for guessed rank at the start of round 1 and its interaction with a dummy for having lost the first-round competition.[17] The gender difference in the reaction to losing for those who initially compete is hardly affected. For the main experiment, the coefficient actually increases in magnitude. However, the gender difference in the reaction to positive feedback

[16] Regressing initial beliefs in the main experiment (predicted probability of winning in round 1) on gender and performance in round 1, the coefficient on the female dummy is −0.081 (0.032) meaning that conditional on performance the average woman thinks that it is around 8 percentage points less likely that she would win a competition against a randomly selected opponent than the average man.

[17] The likelihood of entering competition is not necessarily linear in beliefs. In particular, a risk-neutral person would compete if they think their chance of winning is above 50 percent and would choose the piece rate otherwise. To make the regressions more flexible, we also include a dummy indicating that the predicted chance of winning is above 50 percent and include full interactions of this dummy with the continuous beliefs variable and the competition outcome dummy.

Table 6—Competition Entry after Round 1: The Role of Initial Beliefs, Belief Updating, and Performance in Explaining the Gender Difference in the Reaction to Competition Outcomes

	(1)	(2)	(3)	(4)	(5)
Competition in round 1 (*main experiment*)					
Female × round 1 loser	−0.349	−0.386	−0.354	−0.373	−0.424
	(0.144)	(0.134)	(0.128)	(0.142)	(0.129)
Observations	460	460	460	460	460
Competition in round 1 (*feedback experiment*)					
Female × round 1 loser	−0.273	−0.251	−0.289	−0.300	−0.305
	(0.129)	(0.143)	(0.111)	(0.130)	(0.124)
Observations	261	261	261	261	261
Piece rate in round 1 (*feedback experiment*)					
Female × round 1 winner	−0.287	−0.176	−0.183	−0.253	−0.114
	(0.115)	(0.117)	(0.108)	(0.107)	(0.101)
Observations	291	291	291	291	291
R1 beliefs × Win R1		√			√
Beliefs			√		√
Rank				√	√
R1 score fixed effects	√	√	√	√	√
R1 rank	√	√	√	√	√

Notes: The table shows coefficients from OLS regressions of a binary indicator for choosing the competition in subsequent rounds on a gender dummy, a dummy for having won the round 1 competition, and the interaction of the two. R1 score means score in round 1, R1 rank means normalised within-session rank in round 1, R1 beliefs means probability of winning in round 1 as predicted by the participant, beliefs mean beliefs at the start of each round, and rank means rank in each round. Standard errors are in parentheses. Standard errors are clustered at the individual level. The main experiment consisted of six rounds of three minutes each, and the feedback experiment consisted of four rounds of four minutes each.

shrinks by 39 percent for those who initially choose the piece rate in the feedback experiment.

If women update their beliefs more strongly in response to losing than men, this could likewise explain why women's willingness to compete drops more strongly after a loss than men's willingness to compete. The upper panel of Figure A8 in the online Appendix shows the evolution of beliefs (guessed rank) over the rounds by gender and competition outcome in round 1 for the pooled data from the main and feedback experiments. Beliefs in round 1 are normalized to 0. The figure shows beliefs for participants who choose competition in round 1. We see that, if anything, women update more in the positive direction after winning than men and less in the negative direction after losing than men, which makes it unlikely that differences in belief updating can explain the gender difference in the reaction to losing.[18] The lower panel shows beliefs over the rounds for participants who choose piece rate in round 1 of the feedback experiment. Here, we actually see that the beliefs of male winners increase more over the rounds than the beliefs of female winners. This is further explored in the regressions in column 3 of Table 6, which control

[18] This fits with the results of Möbius et al. (2011) and Buser, Gerhards, and van der Weele (2018), who find that men update beliefs about their own abilities more strongly than women in response to feedback.

for predicted rank at the start of each round.[19] For those who initially compete, the magnitude of the coefficient on the gender interaction actually increases slightly, reflecting that men update their beliefs more strongly in response to losing than women. For those who initially pick the piece rate, on the other hand, the gender difference in the reaction to positive feedback drops by 36 percent.[20]

Past studies have shown that women might perform worse after losing in a competition compared to men in the same situation (Gill and Prowse 2014, Buser 2016). Figure A9 in the online Appendix shows the evolution of relative performance (rank) over the rounds by gender and competition outcome in round 1 for the pooled data from the main and feedback experiments. Rank in round 1 is normalized to 0. The upper panel shows relative performance for participants who choose competition in round 1. We see that, if anything, female winners increase their performance more after winning compared to male winners and decrease their performance less after losing compared to male losers, which makes it unlikely that a differential reaction in performance can explain the gender difference in the reaction to losing.[21] The lower panel shows relative performance over the rounds for participants who choose piece rate in round 1 of the feedback experiment. Again, the observed pattern does not indicate that the performance of women reacts differentially to competition outcomes in a way that could explain the effect on subsequent willingness to compete. In column 4 of Table 6, we control for normalized rank (that is, likelihood of winning) in each round.[22] The interaction coefficient does not change much in any of the regressions and the gender difference in the effect of losing actually increases in both datasets.

Finally, we combine all 3 mechanisms in column 5. It is clear that gender differences in initial beliefs, belief updating, and performance changes cannot explain the gender difference in the reaction to losing for those who initially choose to compete. Conditional on initial beliefs, belief updating, and the evolution of relative performance, the magnitude of the gender-interaction coefficient actually increases in both datasets. On the other hand, these mechanisms explain an important part of the gender difference in the reaction to positive feedback for those who do not compete in round 1 of the feedback experiment, mainly due to the explanatory power of gender differences in initial beliefs and belief updating.

[19] Again, to make the regressions more flexible, we also include a dummy indicating that the predicted chance of winning is above 50 percent and the interaction of this dummy with the continuous beliefs is variable.

[20] Participants who choose to compete are matched with a random other participant chosen among all participants. If participants believe that choosing to compete goes along with a boost in performance, their choice of incentive scheme could also be influenced by their belief about the proportion of others who compete. Note, however, that for such beliefs to explain the observed gender differences, women who initially compete would need to drastically revise upward their beliefs of the likelihood of others to enter the competition between rounds 1 and 2, which seems unlikely.

[21] This result is seemingly inconsistent with the findings of Gill and Prowse (2014) and Buser (2016), but there are some important design differences that make the results hard to compare. First, in both of these studies, all participants are forced to compete, whereas our sample consists of participants who self-selected into the competition. Second, in our design, participants can react to the competition outcome by changing the incentive scheme, whereas in the other studies, participants have no choice on whether to compete or not in the subsequent round.

[22] Again, to make the regressions more flexible, we also include a dummy indicating that the chance of winning is above 50 percent and include full interactions of this dummy with the continuous rank variable and the competition outcome dummy.

We conclude that gender differences in initial beliefs, belief updating, performance changes, and changes in risk preferences cannot explain the gender difference in the reaction to losing. Rather, it seems that losing in a competition has a direct negative impact on the preference for competition of women who are initially willing to compete.

III. Results from the Dutch Math Olympiad

In this section, we will use our field data from the Dutch Math Olympiad to determine whether the gender difference in the reaction to losing carries over to the field. We will start by describing how the Math Olympiad works and why it provides a credible setting to identify the gender difference in the effect of losing a competition on subsequent willingness to compete again.

A. *Background, Data, and Identification Strategy*

The Dutch Math Olympiad is an annual national mathematics competition organized for secondary school students. Its ultimate aim is the selection of a national team to represent the Netherlands in the International Mathematics Olympiad.[23] All Dutch secondary school students from the middle and upper tracks up to grade five are allowed to participate in the first round of the Olympiad.[24] The first round of the competition consists of a two-hour test that is administered in the participants' own schools and graded by their teachers.[25]

Only a predetermined number of best performers in round 1 advance to the second round.[26] Because the threshold score for advancing is flexibly determined to approximately select the N highest performers, the setting resembles a competition in which the top performers win and the rest lose. Participants who are in the fourth year of secondary school or lower are free to participate again the year after regardless of their score. We take the binary decision of whether to compete again one year later as our outcome measure.

We have anonymized data for all participants in the 2010–2014 Olympiads including their score, gender, and whether they participated again the year after. The vast majority of participants are from years four and five of secondary school. We restrict our sample to fourth-year students who are allowed to participate again one year later when they are in their fifth year. Table A3 in the online Appendix shows the total number of participants and the number of participants who were invited to

[23] The International Mathematics Olympiad is the most recognized international maths competition for pre-university students, with around 100 countries participating. Details on the Dutch Math Olympiad can be found at www.wiskundeolympiade.nl.

[24] Secondary education in the Netherlands consists of three different tracks: vocational, general, and pre-university. Because the national final takes place during the following school year, sixth-year students are not allowed to participate.

[25] There is no discretion in grading as the test consists of multiple-choice questions and questions with a clear numerical answer, and points are only awarded for a correct answer.

[26] The number of participants who advance to the second round is fixed in advance. In 2010, the top 700 performers advanced to the second round. To keep up with the increasing number of participants, this number was increased to 800 in 2011 and 1,000 in 2014.

the second round by year. Participation increased significantly over the years covered by our sample from 4,150 in 2010 to 9,161 in 2014. We also show the number of actual participants in the second round. More than 90 percent of those who make the cutoff participate in the second round. The table also shows the corresponding numbers for our sample of fourth-year students and the proportion of female participants in our sample in each year.

Because winning and losing depend on a strict cutoff, the data present us with a sharp regression discontinuity design. Comparing the subsequent participation choices of individuals just below and just above the cutoff makes it possible to estimate the causal effect of losing versus winning on the likelihood of participating again one year later. In particular, we are interested in the gender difference in this effect.

We employ two commonly used approaches to estimating the regression discontinuity design: linear and polynomial (Lee and Lemieux 2010). In the linear regressions, we restrict the sample to observations that are close to the cutoff and control linearly for the forcing variable (first-round score). The polynomial approach consists of using a higher bandwidth (that is, observations further away from the cutoff) and controlling for the forcing variable using a higher order polynomial. In practice, this usually means using a quadratic control (Gelman and Imbens (2014) advise against the use of higher order polynomials and show that confidence intervals based on them have poor properties). The linear approach has the advantage of staying close to the spirit of regression discontinuity by only using observations close to the cutoff. However, in cases where a polynomial is a good approximation of the underlying data, using the polynomial approach is more efficient. The usual recommendation is to use a variety of specifications and bandwidths (Lee and Lemieux 2010, Imbens and Lemieux 2008, Athey and Imbens 2017).

Our design is somewhat unusual as our main interest is not in the discontinuity itself (that is, the effect of losing on the likelihood of participating again a year later), but in the gender difference in the discontinuity (that is, the gender difference in the effect of losing on the likelihood of participating again a year later). In practice, we estimate the following equation:

$$Y_i = \alpha + \delta_1 T_i + \delta_2 F_i + \delta_3 T_i \cdot F_i + f(s) + f(s) \cdot F_i + f(s_i) \cdot T_i + f(s_i) \cdot T_i \cdot F_i + \varepsilon_i,$$

where Y is a binary indicator for participating again a year later, T is a binary indicator for scoring above the cutoff (that is, making the second round), F is a binary indicator for being female, and $f(s)$ is a polynomial function of the amount of points scored s. The parameter of interest is δ_3, which estimates the gender difference in the reaction to losing. We interact the polynomial $f(s)$ with the cutoff indicator T to allow for different slopes left and right of the cutoff, and interact both $f(s)$ and the interaction of $f(s)$ and T with F to allow for different slopes for each gender. We also present regression discontinuity results separately for each gender.

B. *Regression Discontinuity Results*

Our sample consists of 11,591 individuals, which is the universe of fourth-year secondary school students who participated in the Olympiad from 2010 to 2014.

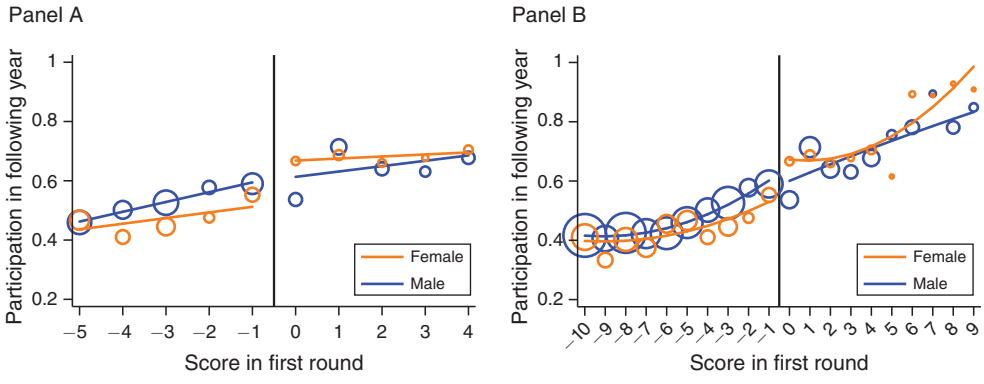

FIGURE 2. REGRESSION DISCONTINUITY GRAPHS

Notes: The *x*-axis shows the score in the first round of the Olympiad. Scores are normalized such that a score of zero or higher means advancing to the second round. The *y*-axis shows the likelihood for participants in a certain bin to participate again in the first round one year later. Panel A shows a scatter plot of observations within a range of 5 points of the cutoff and a linear fit of the data left and right of the cutoff, separately for male and female participants. Panel B shows a scatter plot of observations within a range of 10 points of the cutoff and quadratic fits. The size of the markers is proportional to the amount of observations in that particular bin (score and gender).

Of these, 63.1 percent are male, reflecting the lower willingness to compete of girls in mathematical areas (see, for example, Buser, Niederle, and Oosterbeek 2014); 14.4 percent of male participants and 9.0 percent of female participants score above the cutoff ($p = 0.00$; chi-squared test).[27] Conditional on scoring above the cutoff, the likelihood of actually participating in the second round is virtually identical for male and female participants (91.0 percent versus 89.4 percent; $p = 0.34$). Further, 46.7 percent of all male participants and 41.9 percent of all female participants ($p = 0.00$) participate again the year after.

In Figure 2, we present regression discontinuity graphs using both the linear and polynomial approaches. Panel A shows the data close to the cutoff (plus-minus 5 points) and linear regression lines. The first-round scores are normalized such that a score of zero or above means passing the threshold. The size of the markers is proportional to the number of observations. Panel B shows a wider bandwidth (plus-minus 10 points) and quadratic approximations. In both cases it is evident that there is a sizable drop at the cutoff in the likelihood of participating again for girls but not for boys. Figure A11 in the online Appendix presents the same graphs without added regression lines.

Figures 3 and 4 show OLS estimates of the discontinuity for a range of score bandwidths around the cutoff using the linear and quadratic approaches.[28] We present estimates of the discontinuity separately for boys and girls, as well as estimates of the gender difference in the discontinuity. For the linear approach we start with a

[27] Figure A10 in the online Appendix shows the full distribution of points for boys and girls where we normalized the scores such that a score of 0 or above means passing the threshold.

[28] In all regressions, we follow the recommendation of Lee and Lemieux (2010) for regression discontinuity with a discrete assignment variable and cluster the standard errors at the score level.

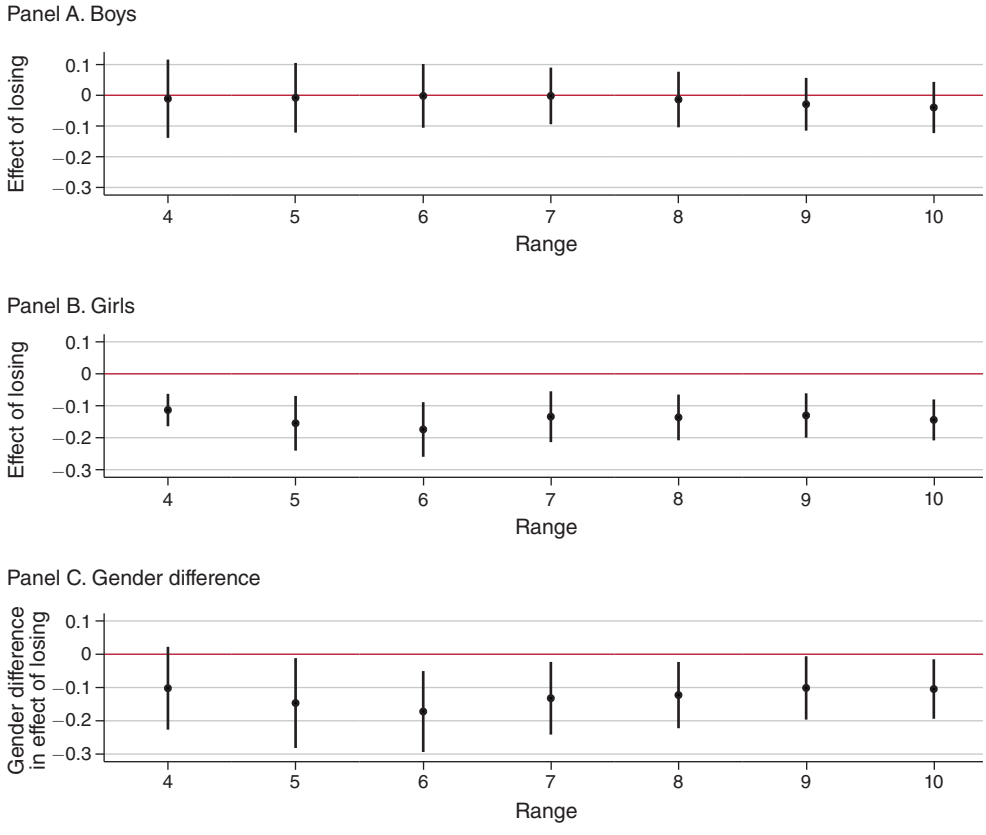

Figure 3. Discontinuity Estimates for Varying Bandwidths (Linear)

Note: Error bars represent 90 percent confidence intervals.

bandwidth of 4 and for the quadratic approach, we start with a bandwidth of 7. The reason is that, because of the very discrete nature of our data (only whole points can be scored), using a second-order polynomial means estimating three parameters either side of the cutoff on a small number of data points. To avoid overfitting, the linear approach is therefore preferable for small bandwidths. To take into account that the likelihood of participating again might also be influenced by a participant's performance relative to his or her gender peers, we also control for within-gender performance percentiles. In Figure A12 in the online Appendix, we present results from regressions without controls for the forcing variable. This approach is essentially a difference-in-differences estimation of the mean likelihood of participating again by gender and first-round outcome. Table A4 in the online Appendix presents the underlying regression coefficients for the difference-in-differences, linear, and quadratic approaches for bandwidths of 4, 8, 11, and 20 points, which roughly corresponds to 25, 50, 75, and 100 percent of the data.

Regardless of the degree of the polynomial and the bandwidth used, the discontinuity estimates for the boys are always very close to zero. The discontinuity estimates for the girls, on the other hand, are consistently between 10 and 20 percentage

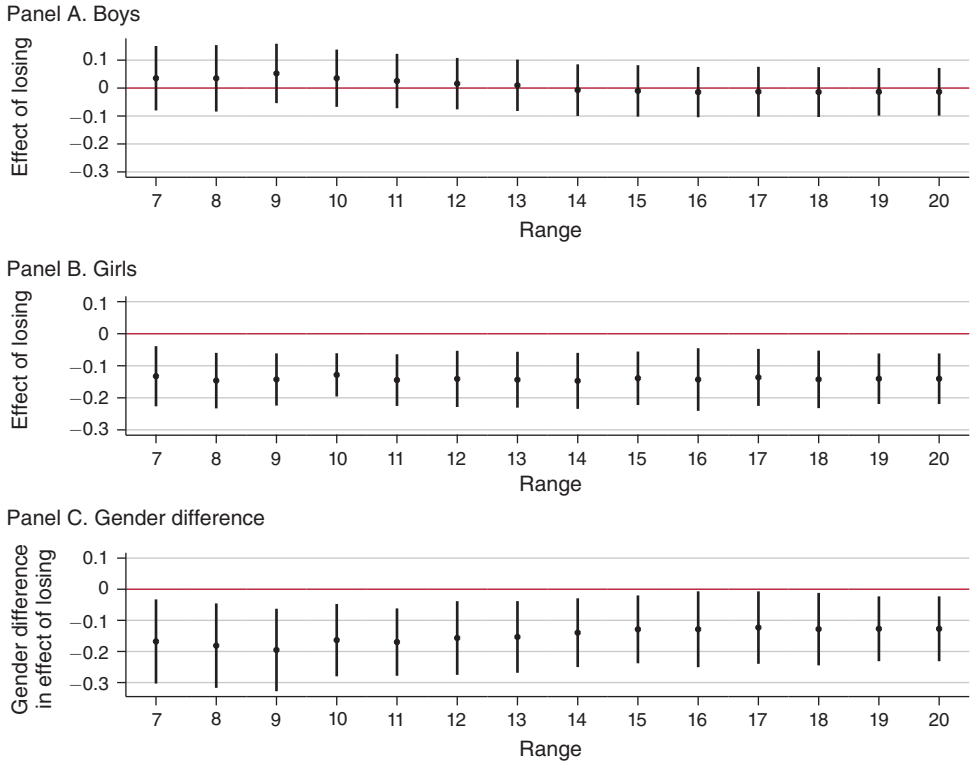

FIGURE 4. DISCONTINUITY ESTIMATES FOR VARYING BANDWIDTHS (QUADRATIC)

Note: Error bars represent 90 percent confidence intervals.

points and always significantly different from 0. Estimates of the gender difference in the discontinuity therefore also mostly fall within the 10 to 20 percentage points range. This means that girls react significantly more strongly to losing than boys. We can compare the discontinuity estimates for the girls to the proportion of girls who score within 5 points of the cutoff who participate again, which is 52 percent. This means that for girls, the effect of losing roughly translates to a 20 to 40 percent reduction in the likelihood of participating again.

These results show that the gender difference in the reaction to losing, which we observed in the lab, carries over to the field. The Math Olympiad provides a relevant setting as past research has shown that willingness to compete in numerical tasks predicts specializing in a science, technology, engineering or mathematics (STEM) related field.[29] Girls being faster to give up on competing in math after a setback could therefore be a partial explanation for the low representation of women in many STEM fields. The regression discontinuity results also demonstrate that the effect of losing persists in the long term. A full year after their first participation, girls who lose are still significantly less likely to participate

[29] See Buser, Niederle, and Oosterbeek (2014) and Buser, Peter, and Wolter (2017a).

again. There are many potential indirect mechanisms for this result. For example, families or teachers could react differently to a negative outcome according to a student's gender. We lack the data to explore this further, but it is important to remember that we find the same gender difference in the lab where such mechanisms are excluded by design.

IV. Conclusions

Differences in willingness to compete between men and women of equal talent are well-documented in the experimental economics literature. Moreover, gender differences in willingness to compete have been shown to partially explain gender differences in career choices and labor market outcomes. This paper asks what happens to the gender gap in willingness to compete in response to winning and losing. We find that men are not only more likely to compete in the first place but they are also more likely to keep competing after losing. This leads to the emergence of a significant gender difference in willingness to compete even among those individuals who start out competing. Using field data from the Dutch Math Olympiad, we not only show that the gender difference in the effect of losing on subsequent willingness to compete carries over to the field, but also that the effect persists in the long term. One year after participating in the Math Olympiad, female losers are significantly less likely to enter again, while there is no effect for male losers.

If these findings translate to career settings, they could be relevant for explaining gender differences in career choices and outcomes. In particular, the gender difference in the reaction to losing is a potential explanation for the so-called "leaking pipeline." In European universities, women make up 59 percent of graduates, 46 percent of PhDs, 37 percent of associate professors, and only 20 percent of full professors. In science and engineering, women make up 35 percent of PhD graduates, 23 percent of associate professors, and just 11 percent of full professors (European Commission 2013). The patterns in the world of business resemble those in academia. The higher up the corporate ladder one goes, the smaller the share of women becomes (Bertrand and Hallock 2001) and gender differences in earnings increase with time after graduation (Bertrand, Goldin, and Katz 2010).

Considering the fierce competition for journal space, research funding, and academic positions, persisting in the face of failure is arguably crucial for being successful in academia. Similarly, it is difficult to imagine somebody making it to the top of the corporate hierarchy without losing out to competitors at some point on the way. High performers who are initially willing to compete are exactly the people we would expect to become future full professors and executives. The gender difference in the reaction to losing, which we document in our data, means that experience and feedback are not guaranteed to eliminate the gender gap in willingness to compete. Rather, our results suggest that bad luck, especially early on in a career, could be especially costly for high-performing women. The results also suggest that framing academic assessments as competitions with winners and losers or providing relative feedback such as rankings, rather than emphasizing absolute performance levels, could have effects on the future choices of students that are not gender neutral.

Future research could focus on the mechanisms behind the gender difference in the reaction to losing and investigate mechanisms for mitigating it.[30] The psychology literature on gender differences in the reaction to success and failure gives some hints for potential psychological mechanisms behind the gender difference in the reaction to losing. According to this literature, men tend to attribute success to internal factors (such as talent) and failure to external factors (such as effort or lack thereof), while women tend to do the opposite (Dweck et al. 1978, Ryckman and Peckham 1987). Future research could also use one of several continuous measures of willingness to compete that have recently appeared in the literature (Saccardo, Pietrasz, and Gneezy 2018; Ifcher and Zarghamee 2016; Petrie and Segal 2015), which could increase statistical power for capturing reactions to feedback. It is also worth mentioning that both our lab and field results stem from competition in mathematical tasks, which are often seen as stereotypically male. Some studies have found that the gender gap in willingness to compete is smaller or zero in stereotypically female tasks, and it would therefore be interesting to investigate whether the gender difference in the reaction to competition outcomes also occurs in neutral or stereotypically female tasks.[31] A further interesting direction would be to investigate whether the propensity to give up competing after a loss predicts individual career choices and outcomes.

REFERENCES

Athey, Susan, and Guido W. Imbens. 2017. "The State of Applied Econometrics: Causality and Policy Evaluation." *Journal of Economic Perspectives* 31 (2): 3–32.

Balafoutas, Loukas, and Matthias Sutter. 2012. "Affirmative Action Policies Promote Women and Do Not Harm Efficiency in the Laboratory." *Science* 335 (6068): 579–82.

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, Armando Jose Garcia Pires, and Bertil Tungodden. 2015. "Competitive in the Lab, Successful in the Field?" *Journal of Economic Behavior and Organization* 118: 303–17.

Berlin, Noémi, and Marie-Pierre Dargnies. 2016. "Gender Differences in Reactions to Feedback and Willingness to Compete." *Journal of Economic Behavior and Organization* 130: 320–36.

Bertrand, Marianne. 2011. "New Perspectives on Gender." In *Handbook of Labor Economics*, Vol. 4B, edited by David Card and Orley Ashenfelter, 1543–90. Amsterdam: North-Holland.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2 (3): 228–55.

Bertrand, Marianne, and Kevin F. Hallock. 2001. "The Gender Gap in Top Corporate Jobs." *Industrial and Labor Relations Review* 55 (1): 3–21.

Boschini, Anne, Anna Dreber, Emma von Essen, Astri Muren, and Eva Ranehill. 2014. "Gender and Economic Preferences in a Large Random Sample." https://ideas.repec.org/p/hhs/sunrpe/2014_0006.html.

Buser, Thomas. 2016. "The Impact of Losing in a Competition on the Willingness to Seek Further Challenges." *Management Science* 62 (12): 3439–49.

Buser, Thomas, Lydia Geijtenbeek, and Erik Plug. 2018. "Sexual Orientation, Competitiveness and Income." *Journal of Economic Behavior and Organization* 151: 191–98.

Buser, Thomas, Leonie Gerhards, and Joël van der Weele. 2018. "Responsiveness to Feedback as a Personal Trait." *Journal of Risk and Uncertainty* 56 (2): 165–92.

[30] See Niederle, Segal, and Vesterlund (2013) and Balafoutas and Sutter (2012) for affirmative action policies that mitigate gender differences in willingness to compete in a static setting.

[31] See Große and Riener (2010); Shurchkov (2012); Dreber, von Essen, and Ranehill (2014); and Boschini et al. (2014). Others, however, have found a similar gap in neutral or stereotypically female tasks (Cárdenas et al. 2012; Wozniak, Harbaugh, and Mayr 2014; Sutter and Glätzle-Rützler 2015; Buser, Peter, and Wolter 2017b).

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, Competitiveness, and Career Choices." *Quarterly Journal of Economics* 129 (3): 1409–47.

Buser, Thomas, Noemi Peter, and Stefan C. Wolter. 2017a. "Gender, Competitiveness, and Study Choices in High School: Evidence from Switzerland." *American Economic Review* 107 (5): 125–30.

Buser, Thomas, Noemi Peter, and Stefan C. Wolter. 2017b. "Gender, Willingness to Compete and Career Choices along the Whole Ability Distribution." IZA Discussion Paper 10976.

Buser, Thomas, and Huaiping Yuan. 2019. "Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad: Dataset." *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20170160.

Cai, Xiqian, Yi Lu, Jessica Pan, and Songfa Zhong. Forthcoming. "Gender Gap under Pressure: Evidence from China's National College Entrance Examination." *Review of Economics and Statistics*.

Cárdenas, Juan-Camilo, Anna Dreber, Emma von Essen, and Eva Ranehill. 2012. "Gender Differences in Competitiveness and Risk Taking: Comparing Children in Colombia and Sweden." *Journal of Economic Behavior and Organization* 83 (1): 11–23.

Cason, Timothy N., William A. Masters, and Roman M. Sheremeta. 2010. "Entry into Winner-Take-All and Proportional-Prize Contests: An Experimental Study." *Journal of Public Economics* 94 (9–10): 604–11.

Charness, Gary, and Uri Gneezy. 2012. "Strong Evidence for Gender Differences in Risk Taking." *Journal of Economic Behavior and Organization* 83 (1): 50–58.

Coutts, Alexander. Forthcoming. "Good News and Bad News Are Still News: Experimental Evidence on Belief Updating." *Experimental Economics*.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47 (2): 448–74.

De Paola, Maria, and Vincenzo Scoppa. 2017. "Gender Differences in Reaction to Psychological Pressure: Evidence from Tennis Players." *European Journal of Work and Organizational Psychology* 26 (3): 444–56.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences." *Journal of the European Economic Association* 9 (3): 522–50.

Dreber, Anna, Emma von Essen, and Eva Ranehill. 2014. "Gender and Competition in Adolescence: Task Matters." *Experimental Economics* 17 (1): 154–72.

Dweck, Carol S., William Davidson, Sharon Nelson, and Bradley Enna. 1978. "Sex Differences in Learned Helplessness: II. The Contingencies of Evaluative Feedback in the Classroom and III. An Experimental Analysis." *Developmental Psychology* 14 (3): 268–76.

Ellis, Jessica, Bailey K. Fosdick, and Chris Rasmussen. 2016. "Women 1.5 Times More Likely to Leave STEM Pipeline after Calculus Compared to Men: Lack of Mathematical Confidence a Potential Culprit." *PLOS One* 11 (7): e0157447.

Ertac, Seda, and Balazs Szentes. 2011. "The Effect of Information on Gender Differences in Competitiveness: Experimental Evidence." https://ideas.repec.org/p/koc/wpaper/1104.html.

European Commission. 2013. *She Figures 2012: Gender in Research and Innovation*. Luxembourg: European Union.

Felder, Richard M., Gary N. Felder, Meredith Mauney, Charles E. Hamrin, Jr., and E. Jacquelin Dietz. 1995. "A Longitudinal Study of Engineering Student Performance and Retention. III. Gender Differences in Student Performance and Attitudes." *Journal of Engineering Education* 84 (2): 151–63.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2): 171–78.

Fischer, Stefanie. 2017. "The Downside of Good Peers: How Classroom Composition Differentially Affects Men's and Women's STEM Persistence." *Labour Economics* 46: 211–26.

Flory, Jeffrey A., Andreas Leibbrandt, and John A. List. 2015. "Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions." *Review of Economic Studies* 82 (1): 122–55.

Gelman, Andrew, and Guido Imbens. 2014. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." NBER Working Paper 20405.

Gill, David, and Victoria Prowse. 2014. "Gender Differences and Dynamics in Competition: The Role of Luck." *Quantitative Economics* 5 (2): 351–76.

Gillen, Ben, Erik Snowberg, and Leeat Yariv. Forthcoming. "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." *Journal of Political Economy*.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics* 118 (3): 1049–74.

Goldin, Claudia. 2013. "Gender and the Undergraduate Economics Major: Notes on the Undergraduate Economics Major at a Highly Selective Liberal Arts College." https://scholar.harvard.edu/files/goldin/files/claudia_gender_paper.pdf.

Goodman, Irene F., Christine M. Cunningham, Cathy Lachapelle, Meredith Thompson, Katherine Bittinger, Robert T. Brennan, and Mario Delci. 2002. *Final Report of the Women's Experiences in College Engineering* (*WECE*) *Project*. Goodman Research Group. Cambridge, MA, April.

Große, Niels Daniel, and Gerhard Riener. 2010. "Explaining Gender Differences in Competitiveness: Gender-Task Stereotypes." Jena Economic Research Paper 2010,017.

▶ **Ifcher, John, and Homa Zarghamee.** 2016. "Pricing Competition: A New Laboratory Measure of Gender Differences in the Willingness to Compete." *Experimental Economics* 19 (3): 642–62.

▶ **Imbens, Guido W., and Thomas Lemieux.** 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142 (2): 615–35.

Iriberri, Nagore, and Pedro Rey-Biel. Forthcoming. "Competitive Pressure Widens the Gender Gap in Performance: Evidence from a Two-Stage Competition in Mathematics." *Economic Journal*.

▶ **Katz, Sandra, David Allbritton, John Aronis, Christine Wilson, and Mary Lou Soffa.** 2006. "Gender, Achievement, and Persistence in an Undergraduate Computer Science Program." *Data Base for Advances in Information Systems* 37 (4): 42–57.

▶ **Lee, David S., and Thomas Lemieux.** 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355.

Legge, Stefan, and Lukas Schmid. 2013. "Rankings, Random Successes, and Individual Performance." Universität St. Gallen Discussion Paper 2013-40.

Möbius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya S. Rosenblat. 2011. "Managing Self-Confidence: Theory and Experimental Evidence." NBER Working Paper 17014.

Niederle, Muriel. 2016. "Gender." In *Handbook of Experimental Economics*, Vol. 2, edited by John Kagel and Alvin E. Roth, 481–562. Berlin: de Gruyter.

▶ **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.

▶ **Niederle, Muriel, and Lise Vesterlund.** 2011. "Gender and Competition." *Annual Review of Economics* 3 (1): 601–30.

▶ **Niederle, Muriel, Carmit Segal, and Lise Vesterlund.** 2013. "How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness." *Management Science* 59 (1): 1–16.

Petrie, Ragan, and Carmit Segal. 2015. "Gender Differences in Competitiveness: The Role of Prizes." George Mason University (GMU) Working Paper in Economics 14-47.

▶ **Price, Joshua.** 2010. "The Effect of Instructor Race and Gender on Student Persistence in STEM Fields." *Economics of Education Review* 29 (6): 901–10.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2015. "Taste for Competition and the Gender Gap among Young Business Professionals." NBER Working Paper 21695.

▶ **Reuben, Ernesto, Matthew Wiswall, and Basit Zafar.** 2017. "Preferences and Biases in Educational Choices and Labor Market Expectations: Shrinking the Black Box of Gender." *Economic Journal* 127 (604): 2153–86.

▶ **Rosenqvist, Olof.** 2019. "Are the Most Competitive Men More Resilient to Failures than the Most Competitive Women? Evidence from Professional Golf Tournaments." *Social Science Quarterly* 100 (3): 578–91.

▶ **Ryckman, David B., and Percy Peckham.** 1987. "Gender Differences in Attributions for Success and Failure Situations across Subject Areas." *Journal of Early Adolescence* 7 (1): 47–63.

▶ **Saccardo, Silvia, Aniela Pietrasz, and Uri Gneezy.** 2018. "On the Size of the Gender Difference in Competitiveness." *Management Science* 64 (4): 1541–54.

Samek, Anya Savikhin. 2015. "Gender Differences in Job Entry Decisions: A University-Wide Field Experiment." Becker Friedman Institute for Research in Economics Working Paper 2579257.

▶ **Shurchkov, Olga.** 2012. "Under Pressure: Gender Differences in Output Quality and Quantity under Competition and Time Constraints." *Journal of the European Economic Association* 10 (5): 1189–1213.

▶ **Sutter, Matthias, and Daniela Glätzle-Rützler.** 2015. "Gender Differences in the Willingness to Compete Emerge Early in Life and Persist." *Management Science* 61 (10): 2339–54.

van Veldhuizen, Roel. 2016. "Gender Differences in Tournament Choices: Risk Preferences, Overconfidence or Competitiveness?" Wissenschaftszentrum Berlin für Sozialforschung Discussion Paper SP II 2016-207.

▶ **Wozniak, David.** 2012. "Gender Differences in a Market with Relative Performance Feedback: Professional Tennis Players." *Journal of Economic Behavior and Organization* 83 (1): 158–71.

Wozniak, David, William T. Harbaugh, and Ulrich Mayr. 2014. "The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices." *Journal of Labor Economics* 32 (1): 161–98.

Zhang, Y. Jane. 2012. "Can Experimental Economics Explain Competitive Behavior Outside the Lab?" http://janezhang.people.ust.hk/assets/files/Zhang_ExtVal.pdf.