Accepted Manuscript

The gender of opponents: Explaining gender differences in performance and risk-taking?

Michael Jetter, Jay K. Walker

PII: S0014-2921(17)30092-2
DOI: 10.1016/j.euroecorev.2017.05.006
Reference: EER 2998

To appear in: *European Economic Review*

Received date: 5 December 2016
Accepted date: 19 May 2017

Please cite this article as: Michael Jetter, Jay K. Walker, The gender of opponents: Explaining gender differences in performance and risk-taking?, *European Economic Review* (2017), doi: 10.1016/j.euroecorev.2017.05.006

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

The gender of opponents:

Explaining gender differences in performance and risk-taking?*

Michael Jetter†   Jay K. Walker‡

May 31, 2017

Abstract

Analyzing 8,169 contestants over 32 years of the US game show *Jeopardy!*, we find that women compete more aggressively, become (marginally) more competitive, and take on more risk when paired against men. Specifically, a woman is more likely to win a *Jeopardy!* episode and, within the show, to respond to a clue. Once responding, she is also marginally more likely to respond correctly. Potentially most surprising, even the sizeable gender gap in risk-taking (analyzing *Daily Double* wagering decisions) disappears once a woman competes in an otherwise all-male field of competitors. Men, on the other hand, wager significantly less when paired against women only, but the gender of opponents does not affect their competitive performance otherwise. Our rich sample allows us to control for a host of potentially confounding factors and player-fixed effects, thereby eliminating potential biases from unobservable individual characteristics. Our findings are consistent with an explanation that emphasizes an adaptation to "social norms" applied to gender.

JEL Classification: D03, D81, G02, J10, J16

Keywords: *competition, financial decision-making, gender differences, performance under high pressure, risk attitudes*

---

*We thank two anonymous referees and the editor Peter Sauer for constructive comments throughout the refereeing process that have helped to substantially improve the paper. We are also grateful to Klaus Abbink, Christian Brown, Christopher Cotton, Nick Feltovich, Lata Gangadharan, Philip Grossman, Wayne Grove, Andrew Hussey, Elias Khalil, Vai-Lam Mui, Andreas Leibbrandt, James Key, Kerry Papps, Patrick Puhani, and Erte Xiao for helpful comments and discussions. We thank seminar participants at Curtin University, Monash University, the University of Central Arkansas, SUNY Buffalo, the University of Memphis, the University of Western Australia, the California State University Fullerton, and the Rochester Institute of Technology for insightful discussions. All remaining errors are our own.

†Corresponding author, University of Western Australia, IZA (Bonn) and CESifo (Munich), 35 Stirling Highway, 6009 Crawley, Australia; Tel: +61 (457) 871-496; Email: mjetter7@gmail.com

‡Department of Economics, Old Dominion University, Constant Hall 2048, Norfolk, VA 23529, USA; E-mail: j1walker@odu.edu

ACCEPTED MANUSCRIPT

1 Introduction

It has almost become a stylized fact that, on average, women are more likely to avoid competition, under-perform in competitive environments, and exhibit higher risk aversion than men. Persistent social phenomena, such as the gender wage gap or the under-representation of women in highly competitive occupations and job positions, have been linked to such observations. One prominent hypothesis to explain this phenomenon relates to the idea that the gender of one's opposition could influence competitive behavior (e.g., see Gneezy et al., 2003, Booth and Nolen, 2012b, Booth et al., 2014). More generally, people may behave differently when competing against adversaries from the opposite sex.

If true, this would imply wide-ranging consequences in a number of settings. For instance, numerous work environments are characterized by persistent under-representation of one gender. Women are especially under-represented in jobs that are generally associated with high-pressure environments and large stakes, such as financial management (the share of females at Wall Street remains at approximately 10 percent) or CEO positions in the US (2.5 percent).[1] Other areas with low female employee shares include IT- and math-related occupations, where women usually occupy less than 20 percent of positions (e.g., see Department of Labor, 2016). On the other hand, men are traditionally under-represented in professions such as nursing, teaching, social work, and counseling. Other prominent settings of gender separation include sports and at times education (e.g., single- versus mixed-gender educational environments).[2]

Unfortunately, real-life settings do not offer many opportunities to test whether the gender of opponents can influence performance and risk-taking since people likely select into gender environments they prefer (consciously or subconsciously) and causality remains difficult to isolate. In the following pages, we study data from the US edition of the popular game show *Jeopardy!* (*J!* from hereon) to analyze whether the exogenous assignment of male or female competitors (i.e., contestants cannot choose to compete against female or male opponents) affects women's

[1] See Eckel and Füllbrunn (2015) for the share of women at Wall Street and Bertrand and Hallock (2001) for the share of women in CEO positions.

[2] For recent discussions about single-gender educational environments, see Jackson (2012) or Park et al. (2013).

ACCEPTED MANUSCRIPT

and men's performance and attitudes to risk. Featuring three contestants per episode, we study competitive behavior in three settings: The probability (i) to win a *J!* episode, (ii) to respond to a given clue, and (iii) to respond correctly to a clue. Finally, to evaluate risk-taking, we exploit the fact that a player can wager up to their entire accumulated balance when faced with a *Daily Double* clue (labeled *DD* from hereon).

A particular strength of our study comes from the large sample of 8,169 contestants in 4,279 *J!* episodes over 32 years, allowing us to control for a rich set of potentially confounding factors, such as scores, clue categories, prior *J!* performance, and player-fixed effects. With fixed effects, we can eliminate the influence of unobserved individual characteristics that may independently influence behavior, including education level, income, age, *J!*-specific knowledge, sense of competitiveness, and risk preferences. The show mimics two important features that are prevalent within leadership positions, such as financial managers or CEOs: Substantial pressure from time constraints and a large viewership, as well as unusually high stakes. On average, *J!* enjoys 25 million viewers per week, and the average winner of an episode earns US$19,752 (see Jeopardy!, 2015d).

When focusing on female contestants, our findings indicate that the same woman is more likely to win an episode and systematically performs more aggressively (i.e., is more likely to respond) when competing against men. For every male counterpart, she is 0.9 percentage points more likely to respond to a given clue. Our estimations also show that the behavior of their male counterparts presents an unlikely explanation of those results. We also find (albeit weaker) evidence of women being more likely to respond correctly to clues when performing in an all-male environment. Further, and maybe most surprisingly, the otherwise robust gender gap in risk-taking disappears once a woman competes in an otherwise all-male environment. In quantitative terms, she wagers 4.9 percentage points more of her maximum possible wager for every male opponent, everything else equal. In turn, the average male *J!* contestant wagers significantly *less* when competing against women. Although the respective magnitude is not as powerful (2.9 percentage points less for every female opponent), that result remains statistically significant at the one percent level.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

The paper proceeds with an overview of the existing literature. Section 3 presents our *J!* data and we refer to the appendix for additional details about *J!*. Section 4 describes the empirical methodology. In Section 5, we describe our results and Section 6 provides a discussion, putting our findings in context to the existing literature and potential policy conclusions. Finally, Section 7 concludes.

2 Background

Gender disparities in (i) the willingness to compete, (ii) performance in competitive situations, and (iii) risk attitudes have been consistently identified in a number of studies.[3] In this Section, we first provide a brief overview of prior work on the determinants of these gender differences, with a particular focus on the gender composition of opponents. We then turn to related studies using game show data that is comparable to our *J!* setting.

2.1 Determinants of Gender Gaps

In recent years, several explanations of gender disparities have been suggested and Table 1 provides a brief overview of that literature. Panel A focuses on competitiveness, whereas Panel B considers attitudes to risk. (Note that we pay less attention to the surrounding literature on the selection into competitive situations.) Proposed drivers include the gender of the opposition (Booth and Nolen, 2012b; De Paola et al., 2015; Säve-Söderbergh and Lindquist, 2016), time constraints (Cotton et al., 2013), but also the form of the task. For instance, verbal versus non-verbal tasks may affect competitive behavior differently for women and men (Shurchkov, 2012). Further, 'gender-specific' tasks may elicit gender differences (Günther et al., 2010; Iriberri and Rey-Biel, 2017) and a team setting may be more conducive to women's performance levels, as

[3]For example, see Gneezy et al. (2003), Gneezy and Rustichini (2004), Niederle and Vesterlund (2007), Fryer and Levitt (2010), Delfgaauw et al. (2013), Ors et al. (2013), Morin (2015), or Heinz et al. (2016) for various settings related to competition. Croson and Gneezy (2009) provide an excellent summary of this literature. Examples for risk attitudes include decisions about smoking and seat belt usage (Hersch, 1996) or financial investment (Jianakoplos and Bernasek, 1998; Eckel and Füllbrunn, 2015). Eckel and Grossman (2008) provide a summary of the experimental evidence on gender differences in risk preferences. A closely related stream of literature considers gender differences in leadership (e.g., see Grossman et al., 2015).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

opposed to an individual setting (Healy and Pate, 2011). In reality, it is likely that explanations are non-exclusive, i.e., gender differences in performance and risk-taking could arise or vanish under a multitude of circumstances or potential combinations thereof. Our focus in this paper will be the gender of opponents and how it may affect women's and men's performance levels and risk-taking in wagering decisions.

This idea related to the opponents' gender has been tested in several experimental studies. For instance, Booth and Nolen (2012a,b) find female teenagers are more likely to *select* into competition and take on more risk in *single*-gender environments (also see Booth et al., 2014). In earlier experimental work, Gneezy et al. (2003) find women to under-perform when paired against males. In turn, De Paola et al. (2015) employ an experimental setup to find no significant performance differences along the lines of opponents' gender for female undergraduate students in Italy.

In general, experimental research designs carry substantial advantages in analyzing this particular question. In reality, people likely select into gender environments they are comfortable with, making it difficult to isolate the causal effect from their opponents' gender amidst various endogeneity concerns. We rarely observe individual performance indicators and expressed risk attitudes of the same woman when *exogenously* assigned to male or female adversaries. Experimental setups can circumvent such selection issues. However, as with any research design, experimental studies also exhibit disadvantages. External validity remains a key concern, as laboratory experiments usually have to rely on relatively small sample sizes.[4] In addition, incentive structures can be somewhat artificial at times. For example, Antonovics et al. (2009) investigate the comparability of results obtained from laboratory settings with data from the field. They find laboratory experiments produce comparable insights only when the offered stakes are high (above US$50) and when players are young (under the age of 33). However, a minimum payoff of US$50 per individual can quickly become expensive if a researcher wishes to draw a large and representative sample. In our context, most experimental studies have found it difficult to cross that payoff threshold (see Table 1).

[4] See Levitt and List (2007) and Niederle (2014) for further discussion on the generalizability of findings from experimental studies.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 1: Selection of recent articles trying to explain the gender gap in performance and risk attitudes (listed chronologically).

Study	Research Setting	Sample	*N*	Payoffs	Main Findings
Panel A: Gender differences in competitiveness					
Iriberri and Rey-Biel (2017)	Experiment	Adults	640	Average 13.80€	Women under-perform when the task if perceived as favoring men
De Paola et al. (2015)	Experiment	Students	720	Exam	Gender of competitors does not affect students' behavior.
Cotton et al. (2013)	Experiment	School kids	505	Candy bar (or small prize) allocated by lottery	Boys outperform girls at first, but in repeated games no gender difference emerges
Shurchkov (2012)	Experiment	Adults	72	Average payoff $44.70	Women perform equally or better in verbal tasks when given more time
Healy and Pate (2011)	Experiment	Adults	192	Average $15 on hourly basis	Women prefer to compete in teams; men prefer individual format
Günther et al. (2010)	Experiment	Students	234	3€ for participation; exact payout varies	Gender differences may depend on tasks at hand (male, female, and gender-neutral tasks)
Gneezy et al. (2003)	Experiment	School kids	324	Approx. US$5 for participation; per unit varies by experiment	Women under-perform in competition, men over-perform; this effect is stronger when women compete against men
Panel B: Gender differences in attitudes to risk					
Säve-Söderbergh and Lindquist (2016)	Game show	Adults & kids	556	Average $1,850 for adults	Women wager less when competing against men
Lindquist and Säve-Söderbergh (2011)	Game show	Adults	615	Max. SEK 88,200 (approx. US$13,569)	Women wager more conservatively if playing against men only
Eckel and Füllbrunn (2015)	Experiment	Adults	108	US$5 for participation; exact payout varies	Men take on more risk than women in investment decisions
Booth et al. (2014)	Experiment	1^{st} year college students	219	Maximum £30	Females studying in single-gender environment are more likely to choose a lottery than females studying in mixed-gender environment
Booth and Nolen (2012b)	Experiment	School kids	260	Average £7	Girls in single-gender schools more likely to select lottery over sure bet than girls in mixed-gender schools
De Roos and Sarafidis (2010)	Real life	Adults	399	A$15,000	Mixed evidence of greater risk aversion by women in Australian *Deal or No Deal* show

Consequently, an important step consists in taking hypotheses from experimental studies to the real world, in order to complement and test conclusions derived in the laboratory. In fact, Levitt and List (2008, p.909) note that "[p]erhaps the greatest challenge facing behavioral economics is demonstrating its applicability in the real world." Related to social norms in particular, Benjamin et al. (2010, p.1913) point out that "it is difficult to test with nonexperimental data whether identity norms play a causal role in economic decision making." Thus, finding a real-life competitive situation where the assignment of female or male opponents cannot be influenced by the contestant could provide us with meaningful information to test the relevance of the gender of opponents.

In this context, *J!* provides a setting that is comparable to real-world competitive situations where pressure is large and decisions have to be made quickly. Lindquist and Säve-Söderbergh (2011) and Säve-Söderbergh and Lindquist (2016) access the Swedish *J!* version to do exactly that. Although their findings indicate no gender gap in risk-taking among children, adult women tend to wager less when competing against men, confirming evidence from Booth and Nolen's (2012b) and Booth et al.'s (2014) experimental studies.

2.2 Related Studies Employing Game Show Data

Recently, several studies have incorporated data from game shows to analyze people's behavior. To our knowledge, Lindquist and Säve-Söderbergh (2011) and Säve-Söderbergh and Lindquist (2016) are the first to focus on how the gender of opponents may influence competitive behavior (also see Table 1), although they are not the first to study behavioral phenomena in a game show environment. Gertner (1993) provides one of the earliest papers to study behavior using game show data, incorporating information from *Card Sharks* to estimate individual risk aversion coefficients for a combined gender sample. Levitt (2004) and Antonovics et al. (2005) use data from the *Weakest Link* to analyze taste and statistical discrimination by race and gender. Post et al. (2008) incorporate data from *Deal or No Deal* to study risk aversion, arguing that reference-dependent theories of decision-making, such as prospect theory, are relevant.

ACCEPTED MANUSCRIPT

Eventually, with large data sets becoming available, researchers have started using game show data to look at potential gender differences. De Roos and Sarafidis (2010) use data from the Australian version of *Deal or No Deal* to look at decision-making under risk, finding mixed evidence of increased risk aversion by females. Kelley and Lemke (2015) employ data from the game show *Cash Cab* to study gender differences in decision-making under uncertainty. They find men to be more likely to accept an end-of-game gamble, but males and females weigh variables differently when using subjective probabilities.

Turning to studies incorporating *J!* data, Metrick (1995) analyzes 1,150 *Final Jeopardy!* round decisions from 1989 to 1992. Much like most of the studies mentioned above, he uses game show data to estimate risk preferences. However, his focus is not on gender differences or the gender composition of opponents. He et al. (2008) study *Daily Double* wagering decisions, finding females to wager less than males, even after having answered questions correctly before in that same category. Although their study addresses potential gender differences in risk preferences, they also do not distinguish by the gender of opponents.

In sum, there exists a well established history of using game show data to study economic decision-making, both in general and in the context of potential gender differences. Nevertheless, the gender of opponents has received less attention as a potential driver of competitive behavior and risk-taking, with the exception of Lindquist and Säve-Söderbergh (2011) and Säve-Söderbergh and Lindquist (2016). Section 6 puts our findings in context to these prior works.

3 *Jeopardy!* Background and Data

3.1 Show Description

On September 10, 1984, *J!* started its ongoing run on television. Each episode hosts three contestants and three rounds (*Jeopardy!*, *Double Jeopardy!*, and *Final Jeopardy!*) with a combined number of 61 clues. Contrary to most other game shows, the host (Alex Trebek) announces 'clues' and the first contestant to respond after the prompt has to pose the correct question.

ACCEPTED MANUSCRIPT

Throughout the paper, we will use the terminology of 'responding' or 'answering' to clues to facilitate readability. For any ordinary clue, the first one to respond correctly receives the associated monetary value toward their balance and is allowed to select the next clue. However, if a contestant is incorrect in their response they will have the dollar value of the clue subtracted from their account balance and the other contestants have the opportunity to respond. The contestant with the highest score at the end of the episode receives their account balance as prize money and is able to return for the next episode, while the second and third place contestants receive consolation prizes.

Each episode begins with the *J!* round consisting of 30 clues (six categories with five clues each), followed by the *Double Jeopardy!* (*DJ!* from hereon) round that also features 30 clues. Note that sometimes not all 30 clues of a given round are completed due to time constraints. The only difference between the *DJ!* and *J!* rounds lies in the fact that all clue values are doubled and clues generally become more difficult (Trebeck and Barsocchini, 1990). Until November 26, 2001, the *J!* round consisted of the clue values \$100, \$200, \$300, \$400, and \$500, whereas the *DJ!* round featured the values \$200, \$400, \$600, \$800, and \$1,000. Since then, all clue values have been doubled.

Eventually, an episode finishes with the *Final Jeopardy!* (*FJ!* from hereon) round that includes a single clue in which each contestant can wager up to their entire account balance on responding correctly. Note that we exclude *FJ!* clues in the main part of the paper, as all contestants naturally respond to that clue and wagering decisions in the *FJ!* round can follow a number of game-theoretic strategies, depending on the account balances of all contestants (see Metrick, 1995). Nevertheless, we will consider the performance in the *FJ!* round when exploring a potential explanation of our main findings in section 5.5.

3.2 Sample Description & Settings

As of June 5, 2015, the *J! Archive* website, a fan-created archive of *J!* episodes, contains full information for 4,279 complete episodes, including 8,169 players (3,726 women and 4,443 men) and 254,079 clues. This includes 12,616 *DD* clues taken by 6,076 contestants throughout the

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

sample. In terms of monetary stakes, the average winner of an episode in our sample takes home US$19,752, equivalent to about four to five times the monthly income of a median household in the United States (see DeNavas-Walt and Proctor, 2015). This compares favorably to Ariely et al.'s (2009) seminal paper on large stakes, where the maximum payoff for Indian workers comes to 400 rupees, which is equivalent to approximately 80 percent of the average monthly consumer expenditure (see Ariely et al., 2009, p.454).

Employing a data specialist, we extracted all information of all available episodes. While the dataset is extensive, it does not include all *J!* episodes. Since January 5, 2004, all episodes are present, but earlier seasons are missing occasional episodes on the website. However, if an episode is present, all corresponding information is available, with the exception of 11 episodes. where the show description indicates the lack of some clues.[5] After checking the episode numbers and available data, we do not find evidence of systematic omissions from the archive. Nevertheless, limiting the sample to episodes #4,451 to #7,084 (since January 5, 2004), where the data exhibits no gaps, produces results consistent with our findings. The corresponding results are available in the appendix Tables AII – AV.

The website contains information about all three contestants' full names, their accumulated prize money, the category of each clue, the sequence of clues, and the value of each clue. Most importantly, the first name of each contestant allows us to conjecture their gender. In most cases, names are commonly attributable to a gender (e.g., Alison and Alyssa are female; Adam and Frank are male) and in those cases where names could indicate either a female or a male contestant, a Google search for the full name readily produces a picture of the *J!* contestant. The same follows if names are abbreviated. This approach allows us to allocate a binary gender indicator to all 8,169 contestants.

Throughout the analysis, we focus on four distinct competitive scenarios. First, we begin by analyzing the most basic outcome of winning an episode. Second, we consider each contestant's probability to respond to a clue. Thus, for each clue we obtain three data points (one for each contestant), producing a sample of 749,433 observations in this setting. Third, we evaluate the

[5]All results are consistent when discarding those incomplete episodes.

probability of responding correctly, producing a sample of 248,052 observations. Note that if the answer is incorrect opponents can choose whether they wish to answer the clue. Thus, clues could be answered by zero (nobody chooses to respond), one, two, or all three contestants (if given answers continue to be incorrect).

Fourth and final, we turn to risk-taking in wagering decisions in *DD* clues. Throughout each episode, three *DD* clues are hidden and if a contestant happens to select one, they can wager up to their entire account balance on responding correctly. We divide the realized wager by the maximum possible wager to estimate what percentage of the maximum sum is wagered (following Säve-Söderbergh and Lindquist, 2016).[6] Everything else equal, a higher percentage indicates higher stakes and therefore higher risk. All our results are consistent when focusing on the absolute wager (see appendix Table AV). For further detail on *J!* and *DD* clues, we refer the reader to the appendix and Jetter and Walker (2016).

3.3 Descriptive Statistics

To introduce the gender composition of the sample, Figure 1 visualizes the share of women over time in *J!* episodes (left top), as well as the gender composition of shows over time (right top). Note that the share of female contestants has consistently increased since the show's inception in 1984. In fact, since 2006, it is possible to take an online test as a first step toward becoming a *J!* participant, which is believed to have increased the number of women on the show (Jeopardy!, 2015d, also see the appendix for more details).

The top right graph and the left bottom graph in Figure 1 also show that all-female episodes are rare and the most common gender composition features one woman and two men (F-M-M). Thus, although the episode-to-episode assignment of female or male competitors is likely exogenous (i.e., no competitor can select the gender of their opponents), a shortage of female *J!* candidates could theoretically lead to a non-random gender assignment. To minimize the

[6] The *J!* rules state that a player can wager up to their entire account balance or up to the largest dollar value on the current board, whichever value of the two is larger. In the initial *J!* and *DJ!* rounds, the largest amount on the board corresponds to US$500 and US$1,000, respectively. Since November 26, 2001, values for all clues have doubled, raising those values to US$1,000 and US$2,000, respectively.

ACCEPTED MANUSCRIPT

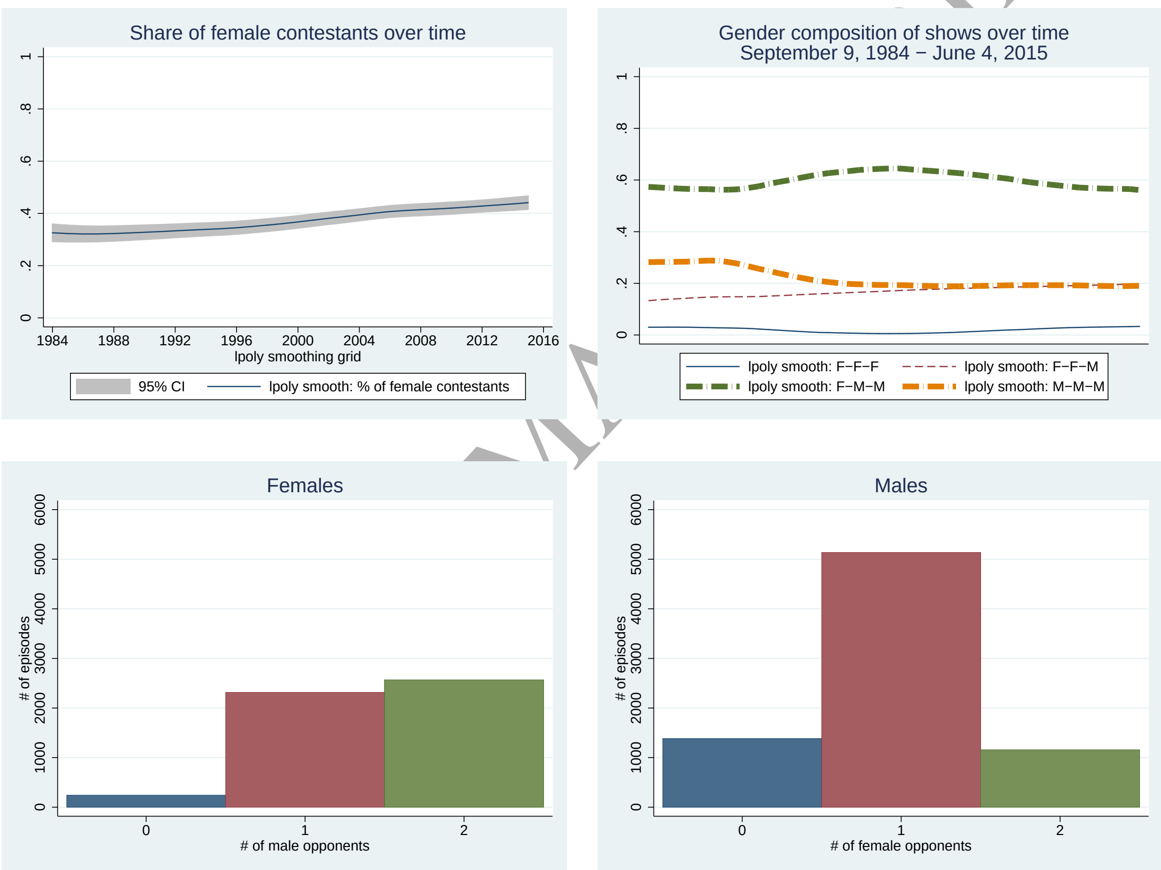

Figure 1: Gender composition of *J!* episodes.

ACCEPTED MANUSCRIPT

influence of such dynamics on our results, we re-estimated all regressions when only using the first episode of the respective contestant. The corresponding results are consistent with our baseline findings and are available in appendix Tables AII – AV.

As a first step toward analyzing contestants' behavior in our four settings of interest, Figure 2 checks for basic gender differences, displaying the respective means with two-sided 95 percent confidence intervals. No statistically meaningful gender differences emerge for the likelihood to win an episode, the likelihood to respond to a given clue, or the likelihood to respond correctly. However, we do observe a pronounced difference in wagering behavior as women usually wager 45 percent of the possible maximum, whereas men wager 48 percent, on average. The difference between these means is statistically significant on the one percent level and is consistent with results from a number of related studies (e.g., Byrnes et al., 1999; Eckel and Füllbrunn, 2015). More detailed descriptive statistics by gender are available in Table AI in the appendix.

To provide a basic idea about whether and how the gender of opponents may influence performance, competitive behavior, and risk-taking, Figures 3 and 4 display means in all four settings of interest when distinguishing by the gender of opponents, along with two-sided 95 percent confidence intervals. In all respective graphs, the horizontal red line provides the mean of the opposite sex as a basic reference point. Focusing on the female perspective, the first three graphs of Figure 3 show that a woman performs better and more aggressively than men (i.e., chooses to respond to more clues) when competing in an otherwise all-male field of competitors. This result is interesting, as not only are women in that setting equal to men, but rather *more* competitive than men. The final graph of Figure 3 (right bottom) shows a similar development for women's wagering behavior. The sizeable gender gap in risk-taking disappears once a woman competes against two males. Note also that, throughout all four settings, the respective confidence intervals become larger for the all-female setting for which relatively few data points exist (only 81 episodes).

Turning to the male perspective, Figure 4 suggests that men perform better in *J!* when competing against two females. They are marginally more likely to win an episode and are more likely to respond to a given clue. Interestingly, men are also marginally less likely to respond

ACCEPTED MANUSCRIPT

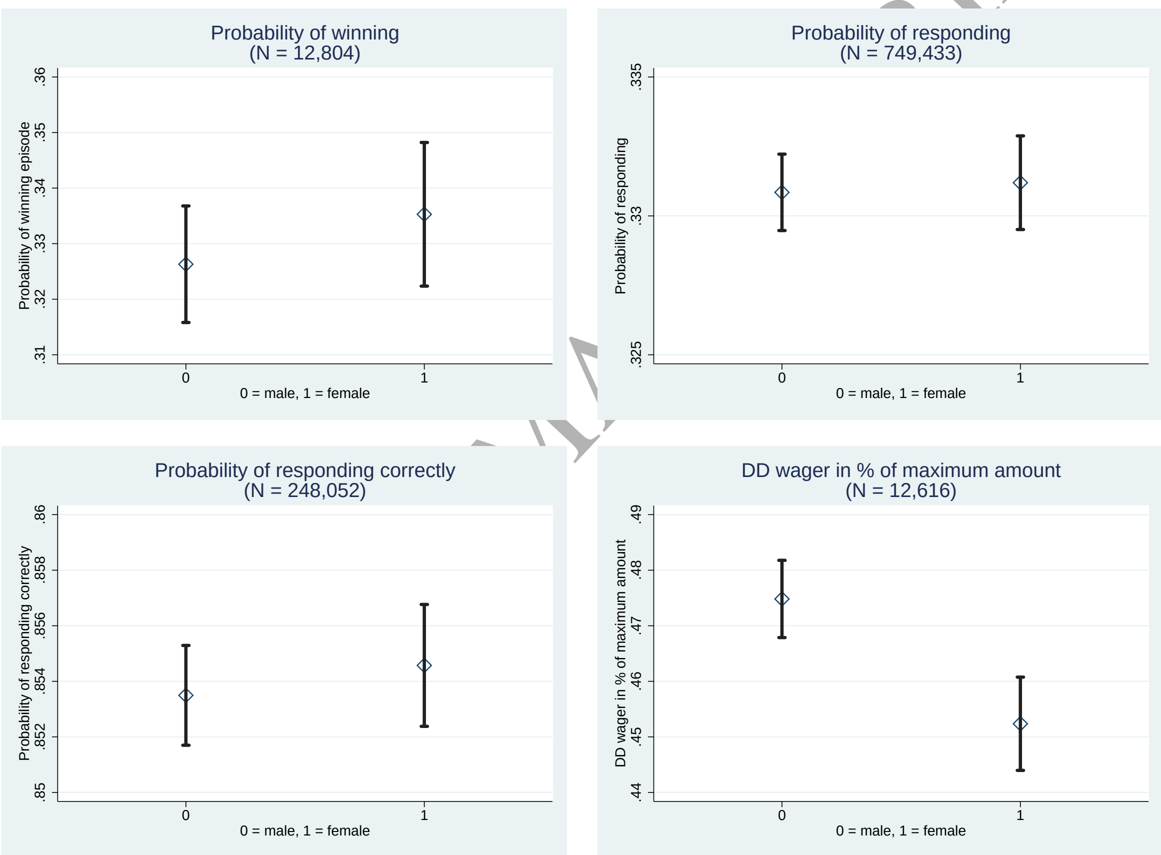

Figure 2: Full sample, comparing females' to males' performance and risk attitudes.

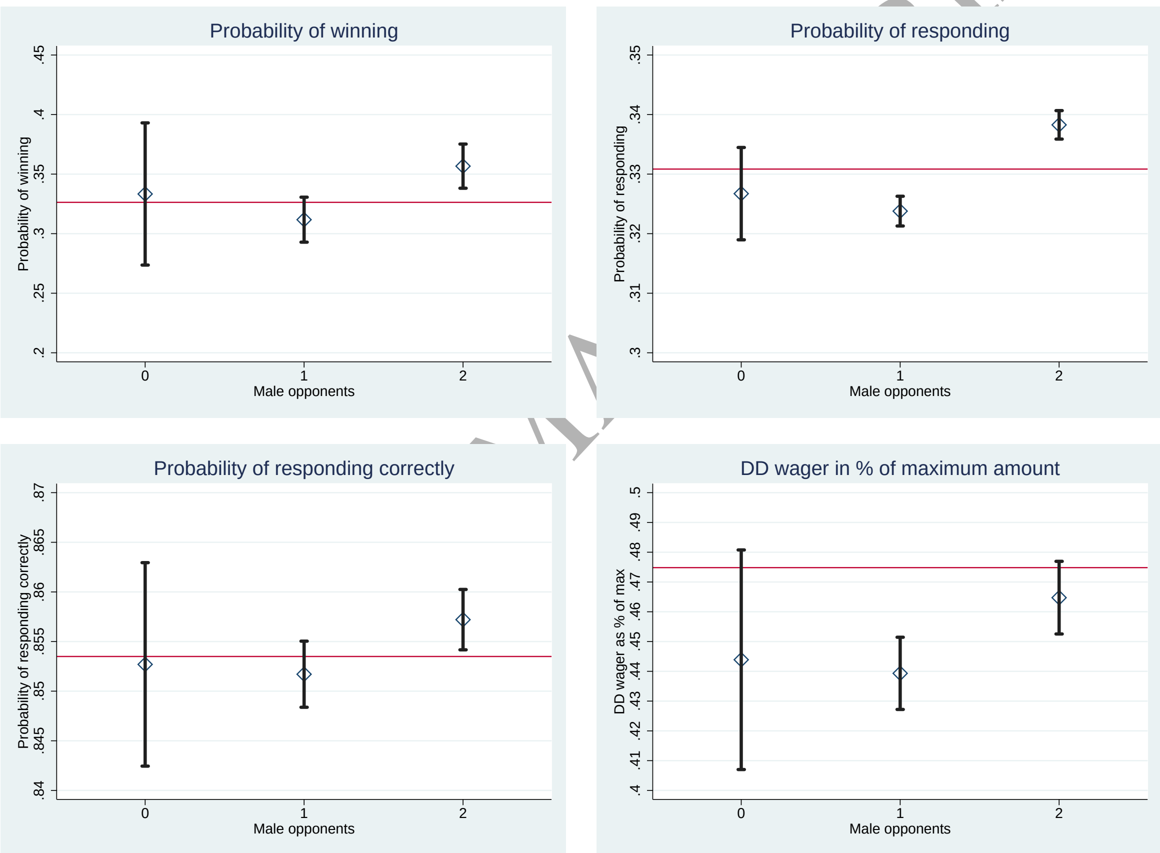

Figure 3: Comparing women's performance and risk attitudes by gender of their opponents.

ACCEPTED MANUSCRIPT

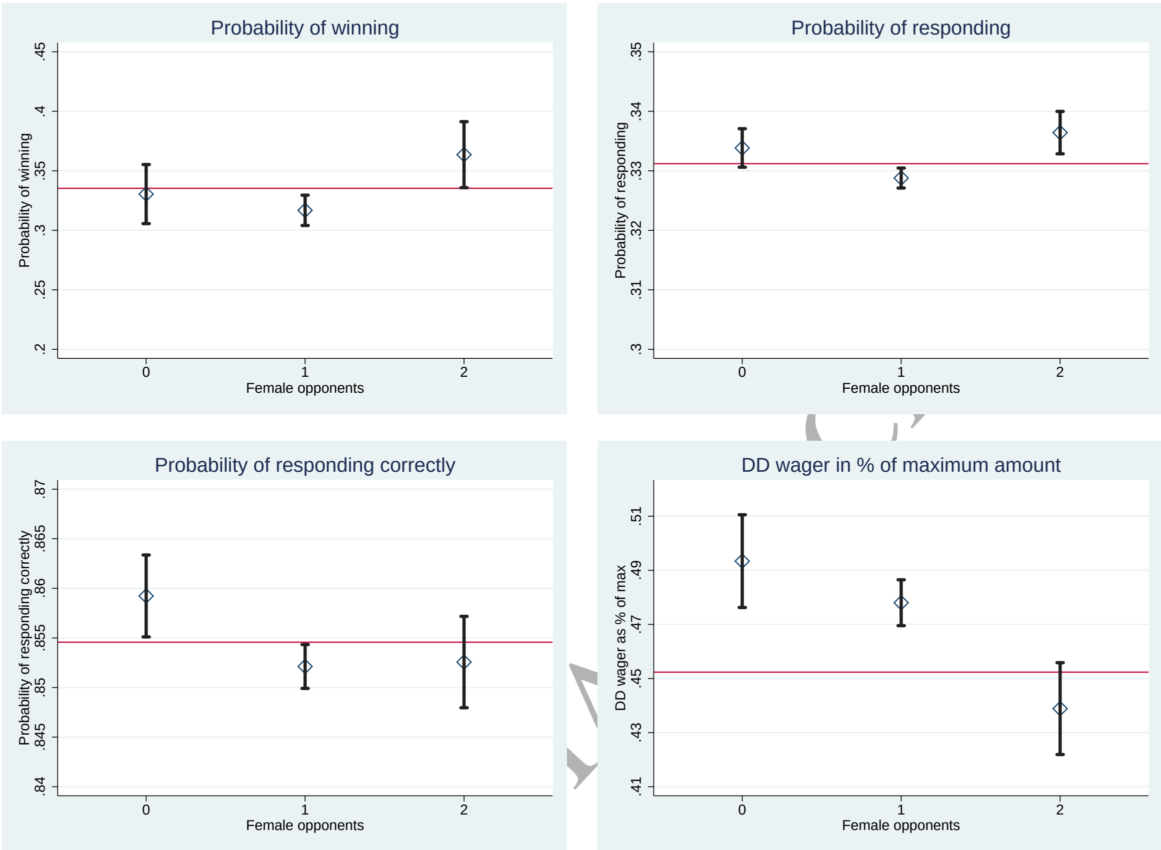

Figure 4: Comparing men's performance and risk attitudes by gender of their opponents.

correctly and wager substantially less when competing against women (bottom graphs of Figure 4). With these basic descriptive statistics in mind, we now turn to describing our empirical strategy.

4 Empirical Strategy

4.1 Methodology

Our econometric strategy follows a standard regression model. We analyze probabilities of winning an episode, responding, and responding correctly via a conventional OLS approach that facilitates the inclusion of player-fixed effects.[7] Nevertheless, the results from employing a logit framework in a pooled setting are virtually identical to those from OLS estimations and the corresponding results are referred to Tables AII – AIV in the appendix. To properly analyze the role of their opponents' gender, we focus on gender-specific samples. Beginning with the female subsample, if $y_{i,t}$ represents one of our four main outcome variables, our estimation becomes

$$y_{i,t} = \alpha_1 \left(male\ opponents\right)_{i,t} + \alpha_2 \boldsymbol{X_{i,t}} + \alpha_3 \boldsymbol{Z_i} + \epsilon_{i,t}. \tag{1}$$

The coefficient associated with α_1 represents the effect of every additional male opponent on female contestant i's performance and her attitudes to risk. The vector $\boldsymbol{X_{i,t}}$ incorporates control variables that could potentially influence $y_{i,t}$. These will be discussed shortly. Further, $\boldsymbol{Z_i}$ corresponds to player-fixed effects, and $\epsilon_{i,t}$ constitutes the conventional error term. Throughout all estimations, errors are clustered at the player level. After focusing on the female subsample, we then re-estimate the same regression analysis displayed in equation 1 for the male subsample, focusing on the number of female opponents as the main regressor of interest.

Concerning $\boldsymbol{X_{i,t}}$, we wish to control for any aspects that may potentially influence the estimation of α_1. In all four settings, we control for the number of *J!* episodes a player has competed in before, as well as year- and month-fixed effects to account for the possibility that (potentially different types of) women and men may have been more likely to participate over the years and in particular times of a given year (e.g., school holidays).[8] For the clue-specific settings, we also incorporate binary indicators for STEM clues (science, technology, engineering, and mathematics) and the 20 most frequent categories on the show, the dollar value of the clue, the account balance of the contestant (both individual and relative to their opponents), the clue number in the current episode, a binary indicator for the *DJ!* round, and an indicator for

[7]By design, fixed effects estimations are difficult to conduct and interpret in a logit or probit framework (Greene, 2004), and the literature then usually moves to the more conventional strategy of employing OLS frameworks.

[8]We thank an anonymous referee for pointing out this possibility. Notice that the number of episodes a player has competed in before is usually equivalent to the number of episodes they have won before. The exception is provided by 'Tournament of Champions' episodes. All results are consistent when measuring this variable as the number of prior wins in *J!* episodes.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

whether the respondent has answered a clue incorrectly previously in this episode.[9] We now briefly explain the intuition behind these control variables.

4.2 Control Variables

In all clue-specific settings, we control for a host of potentially confounding factors. First, research has shown that women tend to opt out of STEM subjects in school and are less likely to pursue STEM-related degrees at University (Preston, 1994; Montmarquette et al., 2002; Griffith, 2010; Grove et al., 2011). Thus, whether a clue can be categorized as STEM may influence performance indicators differently by gender. More generally, it is possible that some categories are preferred by females or males and, in addition to STEM clues, our analysis includes binary indicators for the 20 most common categories to account for such dynamics.

Second, an important line of research focuses on the size of an expected payoff, which directly translates to the dollar value of the clue. A priori, large stakes could lead to increased or decreased performance levels, as the related literature has produced evidence for both. For example, Lazear (2000) suggests higher expected payoffs to raise effort levels and performance. In contrast, Ariely et al. (2009) find that large expected payoffs can lead someone to "choke" and under-perform. Further, the *J!* outline proposes that the degree of difficulty increases with the associated dollar value (Provencher, 2011).

Third, although *J!* participants are "probably all equal in terms of knowledge," according to the host of the show (Trebeck and Barsocchini, 1990, p.61), differences in ability are likely present among the participants. To control for such heterogeneity, we incorporate the current account balance of the respective participant. This variable may not only reflect the contestant's *J!* capabilities, but also their degree of confidence in the ongoing episode. In addition, we capture the relative standing of a player, incorporating a variable relating one's current account balance to their opponents'. This caters to the notion that prior performance of com-

[9]We sorted through all categories, manually sorting them into STEM and non-STEM. The 20 most common categories are science, before & after, literature, potpourri, American history, world history, sports, business & industry, world geography, U.S. cities, colleges & universities, animals, transportation, religion, U.S. geography, opera, authors, people, food, and the Bible. In alternative estimations, we also use time trends (linear and squared) on the daily level, but the corresponding results are virtually identical.

petitors may influence behavior in competitive tasks (Smith, 2013, produces evidence from spelling bee contests). To retain all possible observations, we choose a subtractive formula: $2 \times own\ balance - balance_1 - balance_2$, where subscripts denote opponents.[10] All conclusions are robust to using the difference to the best opponent's balance (positive if leading; negative if trailing) and the corresponding results are referred to Tables AIII – AV in the appendix.

Fourth, we incorporate a variable capturing whether a contestant has responded incorrectly to a clue before in the respective episode. This variable provides another indicator capturing the contestant's *J!* abilities and their confidence level in the ongoing episode. All results are robust when including a binary indicator for whether a contestant has responded correctly before, following results by He et al. (2008) and Post et al. (2008). Fifth, to acknowledge the surrounding characteristics of each clue, we add variables describing the number of the clue in the ongoing round and a binary indicator for the *DJ!* round, where dollar values of all clues are doubled. Both parameters can relate to the accumulated experience of the contestants within the game and previous research has shown that gender differences may disappear in repeated competition (e.g., see Cotton et al., 2013). If such dynamics could indeed affect women's performance differently than men's, then our estimates may be biased if such a variable is not captured.

Sixth, we include year-fixed effects in all five settings. As Figure 1 shows, the share of female contestants has increased over time and we want to isolate our regression results from those underlying developments. In addition, year-fixed effects control for the doubling of prize money since the end of 2001. Finally, month-fixed effects address the idea that *J!* participants from particular professions may select to enter the show at specific times of the year. For example, school teachers – of which a majority are female (Bureau of Labor Statistics, 2016) – may be more likely to enter during school holidays, which could potentially introduce a selection issue into our sample. With this framework in mind, we now turn to our empirical findings.

[10]Note that putting one's score in percentage terms or any other division-based formula would eliminate observations where the denominator is zero. In addition, it could skew observations where the numerator takes on the value of zero. Nevertheless, all derived results are unaffected by choosing such performance indicators.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

5 Empirical Findings

Tables 2 – 5 display regression results for our four dependent variables. In each Table, columns (1) – (3) focus on female *J!* contestants and columns (4) – (6) consider males. In each setup, we first present results from the pooled specification. We then include player-fixed effects in columns (2) and (5), before presenting an alternative version of the full specification where we distinguish between two binary indicators of competing against one or two adversaries from the opposite sex. In any regressions including fixed effects, we only include those contestants who appeared at least in two episodes, since all others would be perfectly identified by their respective dummy variable.

5.1 Setting #1: Winning an Episode

Table 2 focuses on the first and most straightforward setting: Winning a given *J!* episode. In columns (1) and (4), we only control for year- and month-fixed effects, as well as a variable counting the number of *J!* episodes a contestant has competed in before. For women (column 1), the coefficient of interest related to the number of male opponents is positive and statistically significant on the one percent level. In terms of magnitude, each additional male opponent is suggested to raise the chances of winning by 3.4 percentage points. Once player-fixed effects are accounted for (column 2), the size of that coefficient remains stable and, if anything, increases further to 3.8 percentage points. However, standard errors inflate markedly, likely because degrees of freedom decrease once we take into account player-fixed effects. In fact, we only count 2.8 episodes per female contestant, on average, once we rely on returning *J!* winners only in column (2).

Further, column (3) moves from a continuous measure of male opponents to including two binary variables of competing in a single-gender environment (F-F-F) and competing in an otherwise all-male field (F-M-M). The reference category here is determined by episodes with an F-F-M combination. Interestingly, a woman is more likely to win when competing against two males, i.e., when no other woman is present. To place these results in context, assuming the

ACCEPTED MANUSCRIPT

Table 2: Regression results from estimating the likelihood to win a *Jeopardy!* episode.

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Females			Males		
Dependent variable: *Winning episode (0/1, mean = 0.33)*						
# of male opponents	0.0341***	0.0384				
	(0.0117)	(0.0240)				
2 female opponents			0.0369			0.0373
			(0.0581)			(0.0278)
2 male opponents			0.0628**			0.0140
			(0.0286)			(0.0207)
# of female opponents				0.0096	0.0082	
				(0.0099)	(0.0161)	
Control variables[a]	yes	yes	yes	yes	yes	yes
Player-fixed effects		yes	yes		yes	yes
# of players	3,718	782	782	4,432	1,369	1,369
# of episodes	3,807	1,899	1,899	4,187	3,359	3,359
N	5,127	2,191	2,191	7,677	4,614	4,614

Notes: Standard errors clustered on the player level are displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a]Includes year- and month-fixed effects, as well as a variable capturing the number of episodes a player has competed in before (equivalent to the number of episodes they have won).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

winnings of an average *J!* episode, this corresponds to an additional $1,240 in earnings. Turning to the male subsample in columns (4) – (6), we find no discernible link between the gender of opponents and the likelihood of winning an episode. This result emerges when considering the number of female opponents, as well as when introducing dummy variables for two female or two male counterparts in column (6).

5.2 Setting #2: Responding

Turning to the clue-specific settings, we begin by analyzing the probability to respond to a given clue. Note that, in the context of competitive behavior, responding could reflect the choice to respond because the respondent believes they know the answer or it could involve the speed with which a contestant hits the buzzer following the prompt. Both explanations can arguably be considered as behavior in a competitive situation.

Table 3 displays the corresponding results. Column (1) indicates that women are 0.76 percentage points more likely to respond to a given clue for each additional male opponent. This result is statistically significant on the five percent level and remains robust in a player-fixed effects framework (column 2). Further, column (3) shows that competing in an otherwise all-male field of adversaries raises the likelihood of responding by 0.95 percentage points, relative to competing against one female and one male. Thus, women compete more aggressively in *J!* when they face male competitors.[11]

Analyzing the male subsample produces, again, very little heterogeneity along the lines of their opponents' gender. We only find a statistically meaningful distinction in column (6) when contrasting two male opponents against a gender mix of one female and one male opponent. Thus, in an all-male field, men are marginally more likely to respond.[12]

[11] In additional estimations, we also included interaction terms between the binary indicator for two male opponents and the 20 most common clue categories. For instance, it is possible that the sole woman in an otherwise all-male field of contestants carries a 'comparative gender advantage' in some topics and is therefore more likely to respond. However, including such interaction terms does not affect the result displayed in column (3) of Table 3, producing a coefficient of 0.0097 (statistically significant on the ten percent level) for the variable indicating two male opponents.

[12] Note that, for any given clue, between zero and three contestants could respond (if answers continue to be incorrect).

Table 3: Regression results from estimating the likelihood to respond to a clue.

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:		Females			Males	
Dependent variable: *Responding (0/1, mean = 0.331)*						
# of male opponents	0.0076***	0.0094**				
	(0.0020)	(0.0041)				
2 female opponents			-0.0091			0.0057
			(0.0097)			(0.0049)
2 male opponents			0.0095*			0.0084*
			(0.0050)			(0.0045)
# of female opponents				0.0001	-0.0023	
				(0.0017)	(0.0031)	
Control variables[a]	yes	yes	yes	yes	yes	yes
Player-fixed effects		yes	yes		yes	yes
# of players	3,718	782	782	4,432	1,369	1,369
# of episodes	3,807	1,899	1,899	4,187	3,359	3,359
N	299,555	128,390	128,390	449,878	271,298	271,298

Notes: Standard errors clustered on the player level are displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a]Includes binary indicators for STEM clues and the 20 most common categories, the $ value of the clue, the account balance of the contestant (both individual and relative to their opponents), the clue number in the current episode, a binary indicator for the *DJ!* round, an indicator whether the respondent has answered a clue incorrectly before, the number of previous *J!* episodes, and year- and month-fixed effects.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

5.3 Setting #3: Responding Correctly

Moving to the third scenario, Table 4 assesses the probability of responding correctly. This sample considers those 248,052 observations where the respective contestant chose to respond. For women, following the familiar sequence of estimations, the role of their opponents' gender becomes more nuanced than in the previous settings. In the pooled regression displayed in column (1), each additional male opponent raises the likelihood of responding correctly to a given clue by 0.37 percentage points. Note that the statistical precision and magnitude are smaller than in the previous settings, suggesting that the likelihood of winning an episode (setting #1) is only in part driven by a better performance in responding correctly.

In fact, once player-fixed effects are accounted for in column (2), standard errors more than double and, even though the magnitude of the respective coefficient remains virtually identical with 0.0039, this suggests little variation (if any) in women's likelihood of responding correctly by the gender of opponents. Once we distinguish the gender of opponents by binary indicators, the result becomes more pronounced: Competing against two men is associated with an increase of the chance of responding correctly by 0.9 percentage points. Note that this coefficient is derived relative to the baseline gender combination of a woman competing against a mixed gender field.[13]

Finally, we again observe few dynamics for males along the lines of their opponents' gender. Once player-fixed effects are accounted for, both the statistical relevance and magnitude of the derived coefficients suggest no anomalies. With these results related to performance indicators in mind, we now move to analyzing risk-taking in the wagering decisions related to *DD* clues.

5.4 Setting #4: Wagering in Daily Double Clues

Our final setting studies the wagering behavior of contestants who happen to choose a *DD* clue. In Table 5, we follow the same sequence of regressions, this time estimating the respective

[13]Here again, including interaction terms between the binary indicator for two male opponents and each of the 20 main clue categories produces virtually identical results (coefficient of 0.0096, statistically significant on the ten percent level).

ACCEPTED MANUSCRIPT

Table 4: Regression results from estimating the likelihood to respond correctly to a clue.

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:		Females			Males	
Dependent variable: *Responding correctly (0/1, mean = 0.837)*						
# of male opponents	0.0037*	0.0039				
	(0.0020)	(0.0042)				
2 female opponents			0.0122			0.0059
			(0.0103)			(0.0053)
2 male opponents			0.0091*			0.0018
			(0.0054)			(0.0046)
# of female opponents				-0.0035**	0.0015	
				(0.0017)	(0.0032)	
Control variables[a]	yes	yes	yes	yes	yes	yes
Player-fixed effects		yes	yes		yes	yes
# of players	3,718	782	782	4,432	1,369	1,369
# of episodes	3,807	1,899	1,899	4,187	3,359	3,359
N	99,211	42,439	42,439	148,841	86,673	86,673

Notes: Standard errors clustered on the player level are displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a]Includes binary indicators for STEM clues and the 20 most common categories, the $ value of the clue, the account balance of the contestant (both individual and relative to their opponents), the clue number in the current episode, a binary indicator for the *DJ!* round, an indicator whether the respondent has answered a clue incorrectly before, the number of previous *J!* episodes, and year- and month-fixed effects.

ACCEPTED MANUSCRIPT

contestant's wager as a share of their maximum possible wager. (All results are robust to using the absolute wager – see appendix Table AV.)

Table 5: Regression results from estimating the share of the maximum possible wager in *Daily Double* clues.

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:		Females			Males	
Dependent variable: *Share of maximum wager in DD clue (mean = 0.466)*						
# of male opponents	0.0310***	0.0489***				
	(0.0063)	(0.0134)				
2 female opponents			-0.0077			-0.0443***
			(0.0296)			(0.0157)
2 male opponents			0.0623***			0.0171
			(0.0166)			(0.0139)
# of female opponents				-0.0240***	-0.0287***	
				(0.0052)	(0.0099)	
Control variables[a]	yes	yes	yes	yes	yes	yes
Player-fixed effects		yes	yes		yes	yes
# of players	2,683	744	744	3,393	1,297	1,297
# of episodes	2,911	1,392	1,392	3,692	2,513	2,513
N	5,010	2,166	2,166	7,605	4,447	4,447

Notes: Standard errors clustered on the player level are displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a] Includes binary indicators for STEM clues and the 20 most common categories, the $ value of the clue, the account balance of the contestant (both individual and relative to their opponents), the clue number in the current episode, a binary indicator for the *DJ!* round, an indicator whether the respondent has answered a clue incorrectly before, the number of previous *J!* episodes, and year- and month-fixed effects.

Column (1) suggests that every male competitor raises the average woman's wager by 3.1 percentage points. This result is statistically powerful and further *increases* in magnitude to 4.9 percentage points when player-fixed effects are considered in column (2). Thus, the same woman wagers markedly more when competing against men, as opposed to when competing

ACCEPTED MANUSCRIPT

against other women. Column (3) then reveals at which gender combination these dynamics mostly occur. In an otherwise all-male environment, a woman wagers 6.2 percentage points more than when facing one female and one male as her opponents. This result is surprising, especially given prior findings by Lindquist and Säve-Söderbergh (2011) and Säve-Söderbergh and Lindquist (2016), who find women to be more competitive in *single*-gender environments using the Swedish edition of *J!*, as well as evidence from Booth and Nolen (2012b).

When turning to the male subsample, we also find substantial differences along the lines of their opponents' gender. Interestingly, men wager substantially less when competing against women. In this case, the literature has commonly found that men do not respond to the gender of their opponents in terms of risk-taking (e.g., see Booth and Nolen, 2012b, or Croson and Gneezy, 2009). Although the respective magnitude in column (5) remains markedly lower than the one derived from the female subsample (0.0287 versus 0.0489), the effect is estimated precisely in statistical terms. Further, distinguishing by the exact gender composition of their opponents in column (6) shows that a man wagers 4.4 percentage points less when in an otherwise all-female environment.

5.5 One Explanation: Game Strategy

Taken together, these findings are consistent with the idea of people adapting to "social norms" applied to gender roles, which we will describe in more detail in Section 6.1. Although we cannot rule out alternative explanations based on more practical hypotheses in a conclusive manner, we can test for at least one. The argument goes as follows. Within *J!*, some clues are more important than others. In particular, *DD* clues are characterized by much larger stakes than the average clue ($1,922 versus $711). Similarly, the *FJ!* round, in which each contestant can wager up to their entire account balance, arguably represents the most important clue at the end of an episode. In those clues alone, contestants play for an average of $4,841 in our sample. Now, if men performed better in high-stakes situations than women, on average (e.g., see Azmat et al., 2016), then a rational approach might suggest women compete more aggressively throughout the show in order to compensate.

ACCEPTED MANUSCRIPT

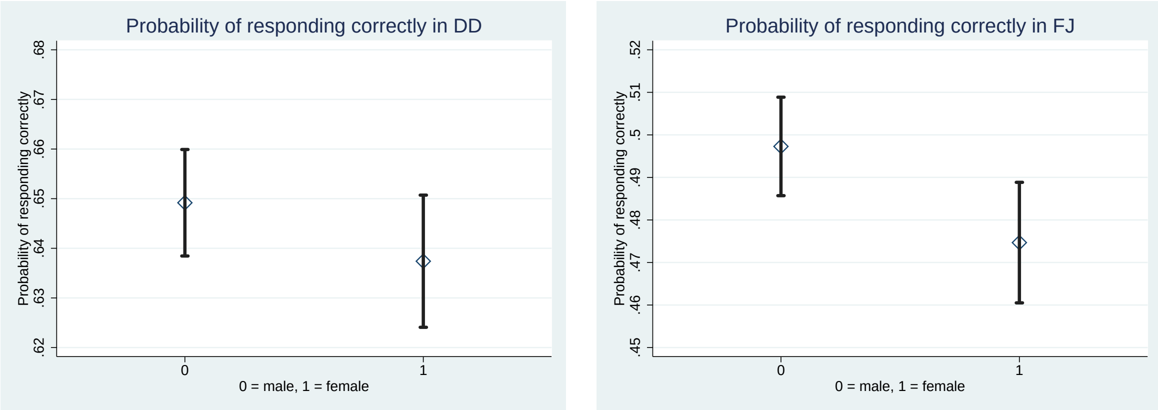

Figure 5: Responding correctly to *DD* (left) and *FJ!* clues (right).

However, analyzing the responses to *DD* and *FJ!* clues produces no discernible gender differences which would support that idea. The results are visualized in Figure 5, again displaying two-sided 95 percent confidence intervals for gender-specific means. Thus, it appears unlikely that women consciously compete more aggressively in regular *J!* clues to compensate for potential gender differences in the performance in high-stakes clues – simply because there are no such gender differences.

6 Discussion

6.1 Interpretation of Results

Our findings are consistent with the idea of adapting to "social norms" applied to the gender of one's opponents. As Datta Gupta et al. (2013) put it, people may use their own and the other participants' gender as a "coordination device." More generally, Akerlof and Kranton (2013, p.11) write that "[t]he norms of how to behave depend on people's positions within their social context."[14] Indeed, both female and male contestants appear to adjust their risk-taking strategy

[14]In recent years, gender norms and identities have been increasingly of note in economics with arguments being placed forth that differences in norms related to social identities can help explain decisions and economic outcomes (Akerlof and Kranton, 2000; Benjamin et al., 2010). As an example pertaining to gender, Bertrand et al. (2015) tie identity to marriage satisfaction and divorce when females have higher relative income to males within households.

in *DD* clues to the gender composition of their opponents: As men generally wager more than women, a woman increases her average wager when more men are present. A man, on the other hand, decreases his average wager as more women are present.

6.2 Comparison to Related Literature

Our results stand in contrast to those from a few existing studies that analyze the influence of the opponents' gender on performance and risk-taking. Most notably, Lindquist and Säve-Söderbergh (2011) and Säve-Söderbergh and Lindquist (2016) find women wager more conservatively when competing against males. Similarly, Booth and Nolen (2012b) conclude girls to be more competitive and to take on higher risks in a *single*-gender environment. They employ an experimental setup with teenagers, randomly grouping them with male or female counterparts. Our results produce the opposite conclusions, employing data from the US version of *J!*. Finally, De Paola et al. (2015) find no significant performance differences for female undergraduate students from Italy. Thus, we believe our paper is the first to suggest women perform better and take on more risk when paired against men.

We can think of three explanations for why our results may, at first glance, contradict findings from related studies. First, cross-country heterogeneity may explain differential findings. For example, Lindquist and Säve-Söderbergh (2011) and Säve-Söderbergh and Lindquist (2016) study the Swedish version of *J!*, whereas Booth and Nolen (2012b) analyze experimental evidence from teenagers in the United Kingdom. However, the dynamics of risk preferences may vary systematically across countries and cultures. For example, Cárdenas et al. (2012) show that Colombian and Swedish schoolgirls exhibit substantial differences in attitudes to risk, and Andersen et al. (2013) find differences in how females compete after puberty between matrilineal and patriarchal societies. Thus, our findings could be complementary to the existing literature, using data from the US, a country of immigrants with a tradition of multiculturalism.

Second, the age of contestants differs across studies, as many related papers evaluate the behavior of children or teenagers. Behavioral differences between children and adults have been highlighted in a number of competitive settings (e.g., see Andersen et al., 2013). Thus, it is well possible – and in fact likely – that competing against adversaries from the opposite sex elicits different reactions from children and teenagers than from adults. This explanation would also suggest a complementarity between our results and some of the existing studies.

Third, our data set substantially extends previous analyses along the lines of sample size and stakes. Our sample includes 8,169 contestants – more than 12 times more than the closest predecessor (Säve-Söderbergh and Lindquist, 2016). For instance, Säve-Söderbergh and Lindquist (2016) are able to incorporate 29 *DD* clues where a woman competes against two men, whereas our US setting includes 2,527 such observations.[15] Thus, more information may simply be more powerful in revealing the underlying general relationship and our paper would serve as contradicting evidence to the discussed studies.

6.3 Policy Relevance

From a policy perspective, our results imply that the gender of opponents can significantly influence behavior and risk-taking in competitive situations. If we want to focus on understanding gender gaps in such behavior, our results suggest women compete more aggressively and risk more when in the company of males. Similarly, men may risk less when competing against women. This result related to risk preferences compares well with that from Eckel and Füllbrunn (2015) who propose that 'excessive' risk-taking in the finance industry could potentially be mitigated by inserting more women in that industry.

To understand the overall degree of performance and risk-taking in competitive fields of different gender compositions, Figure 6 visualizes the overall likelihood of responding correctly (left) and the overall wager in *DD*s (right). Both graphs compare the means of all possible gender compositions, moving from left to right: M-M-M, M-M-F, M-F-F, and F-F-F. The performance average, measured by the likelihood of responding correctly, remains highest in the all-male setting and lowest in the M-F-F combination. Here again, inferences from the all-female setting remain limited because few *J!* episodes have produced that combination.

[15] At this point, it is also useful to highlight the differences in stakes: Although scores are equivalent in the Swedish setting studied by Lindquist and Säve-Söderbergh (2011) and Säve-Söderbergh and Lindquist (2016), $1 translates to approximately 8 SEK. Thus, stakes are substantially higher in the US setting of *Jeopardy!*.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

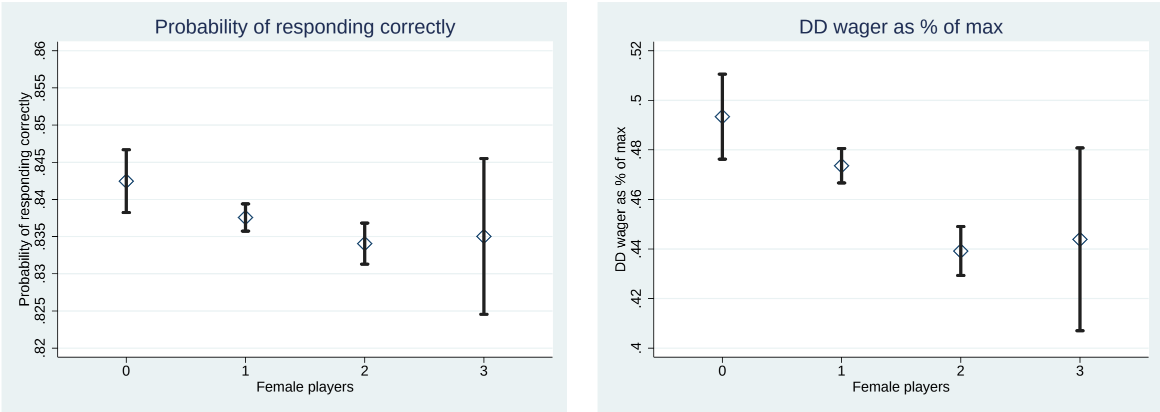

Figure 6: Overall average performance and risk indicators by gender composition of *J!* episodes.

Interestingly, we observe more pronounced differences when assessing risk-taking in *DD* wagers (right graph). Consistent with Eckel and Füllbrunn (2015), the overall degree of risk-taking is highest in an all-male field of competitors, whereas less risk is taken by an M-F-F and an F-F-F field of competitors.[16] Although women take on more risk when in the company of men, the overall degree of risk-taking in *J!* episodes is minimized in the M-F-F setting – likely because men take on substantially less risk when competing against two females (see column 6 of Table 5).

Related to potential policy recommendations, this implies that if we want to lower risk-taking in a given male-dominated industry (say, the finance industry) inserting more women may provide a powerful measure. The same policy recommendation emerges if we are concerned with women's competitive behavior and if society desires women to become more competitive. Nevertheless, one should of course not forget our game show setting and any interpretation should keep the limitations and boundaries of our study in mind. *J!* contestants are not necessarily representative of the entire population, although the show does exhibit the traits we would like to evaluate with respect to performance and attitudes to risk: High stakes and substantial pressure. The fact that our results emerge forcefully when relying on within-player

[16]Relative to Eckel and Füllbrunn (2015), Cueva and Rustichini (2015) argue gender-isolated markets may lead to comparable deviations from fundamentals in asset pricing and mixed gender environments may be more stable.

variation only hints at a general pattern in how the gender of one's opponents can influence performance indicators and attitudes to risk.

7 Conclusion

This paper analyzes a rich database of 4,279 *Jeopardy!* episodes with 8,169 contestants to study whether the gender of one's opponents affects behavior in a highly competitive situation with large stakes. As contestants are unable to choose the gender of their opponents, *Jeopardy!* provides an attractive field setting to explore such dynamics, with individual outcomes being readily available for the researcher.

We study four distinct scenarios: The likelihood (i) to win an episode, (ii) to respond to a clue, (iii) to respond correctly to a clue, and (iv) the wagering decisions in *Daily Double* clues. Contrary to existing studies, we find that a woman is more likely to win and competes more aggressively when paired against males. Further, the otherwise robust gender gap in risk-taking disappears once a woman competes in an all-male field of competitors. These results are robust to the inclusion of a rich list of potentially confounding variables and player-fixed effects, which allow us to control for any unobservable differences on the individual level. Further, these results are unlikely to be driven by a strategic consideration of women performing more aggressively because of a potential under-performance in particularly high-stakes clues. From the male perspective, we find performance indicators to be less responsive to their opponents' gender, but a notable heterogeneity emerges for wagering decisions. In particular, a man wagers significantly less when competing against women.

We discuss potential policy consequences in Section 6, but advise caution in the interpretation and generalization of our results. At the least, we hope that this study stimulates further research on the role of the opponents' gender in influencing performance and risk-taking in high-pressure situations.

ACCEPTED MANUSCRIPT

References

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Akerlof, G. A. and Kranton, R. E. (2013). *Identity economics: how our identities shape our work, wages, and well-being*.

Andersen, S., Ertac, S., Gneezy, U., List, J. A., and Maximiano, S. (2013). Gender, competitiveness, and socialization at a young age: Evidence from a matrilineal and a patriarchal society. *The Review of Economics and Statistics*, 95(4):1438–1443.

Antonovics, K., Arcidiacono, P., and Walsh, R. (2005). Games and discrimination: Lessons from the "Weakest Link". *Journal of Human Resources*, 40(4):918–947.

Antonovics, K., Arcidiacono, P., and Walsh, R. (2009). The effects of gender interactions in the lab and in the field. *The Review of Economics and Statistics*, 91(1):152–162.

Ariely, D., Gneezy, U., Loewenstein, G., and Mazar, N. (2009). Large stakes and big mistakes. *The Review of Economic Studies*, 76(2):451–469.

Azmat, G., Calsamiglia, C., and Iriberri, N. (2016). Gender differences in response to big stakes. *Journal of the European Economic Association*, 14(6):1372–1400.

Benjamin, D. J., Choi, J. J., and Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, 100(4):1913–28.

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

Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, 130(2):571–614.

Booth, A., Cardona-Sosa, L., and Nolen, P. (2014). Gender differences in risk aversion: Do single-sex environments affect their development? *Journal of Economic Behavior & Organization*, 99:126–154.

Booth, A. and Nolen, P. (2012a). Choosing to compete: How different are girls and boys? *Journal of Economic Behavior & Organization*, 81(2):542–555.

Booth, A. L. and Nolen, P. (2012b). Gender differences in risk behaviour: Does nurture matter? *The Economic Journal*, 122(558):F56–F78.

Bureau of Labor Statistics (February 10, 2016). Labor force statistics from the current population survey. https://www.bls.gov/cps/aa2015/cpsaat11.htm. Accessed: 2017-3-25.

ACCEPTED MANUSCRIPT

Byrnes, J. P., Miller, D. C., and Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, 125(3):367.

Cárdenas, J.-C., Dreber, A., Von Essen, E., and Ranehill, E. (2012). Gender differences in competitiveness and risk taking: Comparing children in Colombia and Sweden. *Journal of Economic Behavior & Organization*, 83(1):11–23.

Cotton, C., McIntyre, F., and Price, J. (2013). Gender differences in repeated competition: Evidence from school math contests. *Journal of Economic Behavior & Organization*, 86:52–66.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474.

Cueva, C. and Rustichini, A. (2015). Is financial instability male-driven? Gender and cognitive skills in experimental asset markets. *Journal of Economic Behavior & Organization*, 119:330–344.

Datta Gupta, N., Poulsen, A., and Villeval, M. C. (2013). Gender matching and competitiveness: Experimental evidence. *Economic Inquiry*, 51(1):816–835.

De Paola, M., Gioia, F., and Scoppa, V. (2015). Are females scared of competing with males? Results from a field experiment. *Economics of Education Review*, 48:117–128.

De Roos, N. and Sarafidis, Y. (2010). Decision making under risk in Deal or No Deal. *Journal of Applied Econometrics*, 25(6):987–1027.

Delfgaauw, J., Dur, R., Sol, J., and Verbeke, W. (2013). Tournament incentives in the field: Gender differences in the workplace. *Journal of Labor Economics*, 31(2):305–326.

DeNavas-Walt, C. and Proctor, B. D. (September 2015). Income and poverty in the United States: 2014. www.census.gov/content/dam/Census/library/publications/2015/demo/p60-252.pdf. Accessed: 2016-07-26.

Department of Labor (2016). Computer and information technology occupations. *U.S. Department of Labor, retrieved from* www.dol.gov/wb/stats/Computer-information-technology.htm *on January 13, 2016*.

Eckel, C. C. and Füllbrunn, S. (2015). Thar she blows? Gender, competition, and bubbles in experimental asset markets. *The American Economic Review*, 105(2):906–20.

Eckel, C. C. and Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of Experimental Economics Results*, 1:1061–1073.

Fryer, R. G. and Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, 2(2):210–240.

ACCEPTED MANUSCRIPT

Gertner, R. (1993). Game shows and economic behavior: Risk-taking on "Card Sharks". *The Quarterly Journal of Economics*, 102(2):507–521.

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

Gneezy, U. and Rustichini, A. (2004). Gender and competition at a young age. *American Economic Review Papers & Proceedings*, 94(2):377–381.

Greene, W. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *The Econometrics Journal*, 7(1):98–119.

Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, 29(6):911–922.

Grossman, P. J., Komai, M., and Jensen, J. E. (2015). Leadership and gender in groups: An experiment. *Canadian Journal of Economics/Revue canadienne d'économique*, 48(1):368–388.

Grove, W. A., Hussey, A., and Jetter, M. (2011). The gender pay gap beyond human capital: Heterogeneity in noncognitive skills and in labor market tastes. *Journal of Human Resources*, 46(4):827–874.

Günther, C., Ekinci, N. A., Schwieren, C., and Strobel, M. (2010). Women can't jump? An experiment on competitive attitudes and stereotype threat. *Journal of Economic Behavior & Organization*, 75(3):395–401.

He, X., Inman, J. J., and Mittal, V. (2008). Gender jeopardy in financial risk taking. *Journal of Marketing Research*, 45(4):414–424.

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

Heinz, M., Normann, H.-T., and Rau, H. A. (2016). How competitiveness may cause a gender wage gap: Experimental evidence. *European Economic Review*, 90:336–349.

Hersch, J. (1996). Smoking, seat belts, and other risky consumer decisions: Differences by gender and race. *Managerial and Decision Economics*, 17(5):471–481.

Iriberri, N. and Rey-Biel, P. (2017). Stereotypes are only a threat when beliefs are reinforced: On the sensitivity of gender differences in performance under competition to information provision. *Journal of Economic Behavior & Organization*, 135:99–111.

Jackson, C. K. (2012). Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in Trinidad and Tobago. *Journal of Public Economics*, 96(1):173–187.

ACCEPTED MANUSCRIPT

Jeopardy! (2015a). Contestant FAQs. http://www.jeopardy.com/beacontestant/contestantfaqs/. Accessed: 2015-09-21.

Jeopardy! (2015b). J! archive show #4089 description. http://www.j-archive.com/showgame.php?game_id=1101. Accessed: 2015-12-28.

Jeopardy! (2015c). Jeopardy! full crew listing. https://www.jeopardy.com/about/cast/full-staff. Accessed: 2015-12-29.

Jeopardy! (2015d). Show history. http://www.jeopardy.com/showguide/abouttheshow/showhistory/. Accessed: 2015-09-21.

Jetter, M. and Walker, J. K. (2016). Anchoring in financial decision-making: Evidence from the field. *IZA Discussion Paper # 10151*.

Jianakoplos, N. A. and Bernasek, A. (1998). Are women more risk averse? *Economic Inquiry*, 36(4):620–630.

Kelley, M. R. and Lemke, R. J. (2015). Gender differences when subjective probabilities affect risky decisions: An analysis from the television game show Cash Cab. *Theory and Decision*, 78(1):153–170.

Lazear, E. P. (2000). Performance pay and productivity. *The American Economic Review*, 90(5):1346–1361.

Levitt, S. D. (2004). Testing theories of discrimination: Evidence from "Weakest Link". *Journal of Law and Economics*, 47(2):431–452.

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

Levitt, S. D. and List, J. A. (2008). Homo economicus evolves. *Science*, 319(5865):909–910.

Lindquist, G. S. and Säve-Söderbergh, J. (2011). "Girls will be girls", especially among boys: Risk-taking in the "Daily Double" on *Jeopardy*. *Economics Letters*, 112(2):158–160.

McCown, A. (March 12, 2015). What's it like to be one of the Jeopardy! writers? http://www.avclub.com/article/whats-it-be-one-jeopardy-clue-writers-216093. Accessed: 2015-12-28.

Metrick, A. (1995). A natural experiment in "Jeopardy!". *The American Economic Review*, 85(1):240–53.

Montmarquette, C., Cannings, K., and Mahseredjian, S. (2002). How do young people choose college majors? *Economics of Education Review*, 21(6):543–556.

Morin, L. (2015). Do men and women respond differently to competition? Evidence from a major education reform. *Journal of Labor Economics*, 33(2):443–491.

ACCEPTED MANUSCRIPT

Niederle, M. (2014). Gender. *NBER Working Paper 20788*.

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

Ors, E., Palomino, F., and Peyrache, E. (2013). Performance gender gap: Does competition matter? *Journal of Labor Economics*, 31(3):443–499.

Park, H., Behrman, J. R., and Choi, J. (2013). Causal effects of single-sex schools on college entrance exams and college attendance: Random assignment in Seoul high schools. *Demography*, 50(2):447–469.

Post, T., van den Assem, M. J., Baltussen, G., and Thaler, R. H. (2008). Deal or no deal? Decision making under risk in a large-payoff game show. *The American Economic Review*, 98(1):38–71.

Preston, A. E. (1994). Why have all the women gone? A study of exit of women from the science and engineering professions. *The American Economic Review*, 84(5):1446–1462.

Provencher, K. (2011). Tufts now: The best job for a trivia-loving English major. `http://now.tufts.edu/articles/best-job-trivia-loving-english-major`. Accessed: 2015-12-28.

Säve-Söderbergh, J. and Lindquist, G. S. (2016). Children do not behave like adults: Gender gaps in performance and risktaking in a random social context in the high-stakes game shows Jeopardy and Junior Jeopardy. *Forthcoming in The Economic Journal*.

Shurchkov, O. (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.

Smith, J. (2013). Peers, pressure, and performance at the national spelling bee. *Journal of Human Resources*, 48(2):265–285.

Trebeck, A. and Barsocchini, P. (1990). *The Jeopardy Book*. New York: Harper Collins.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Appendix

Jeopardy!

Every *Jeopardy!* episode consists of three rounds with a total of 61 clues: The *Jeopardy!* (30 clues), *Double Jeopardy!* (30 clues), and *Final Jeopardy!* rounds (1 clue). The *Jeopardy!* round contains six categories of five clues each with values of US$200, 400, 600, 800, and 1,000. After that, the *Double Jeopardy!* round includes the same number of clues (new categories), but twice the prize money.[17] In each *Jeopardy!* round, one *Daily Double* clue is hidden, whereas two are included in the *Double Jeopardy!* round. If a contestant happens to select a *Daily Double* clue, they are allowed to wager up to their entire account balance or the largest dollar value on the current board, whichever of the two is larger. If they respond correctly, they will gain the wagered sum, whereas they will lose that sum if their response is incorrect.

From a 2015 interview with the head *Jeopardy!* writer, a *Daily Double* is "something that requires a little more thought, requires a two-step process" relative to the other clues in the *Jeopardy!* and *Double Jeopardy!* rounds (McCown, 2015). The *Final Jeopardy!* round consists of a single clue in which all contestants can wager up to the total amount of their current account before the clue is given, although the category is known in advance (Trebeck and Barsocchini, 1990, p.171-174). The goal of the game is to finish the three rounds with the most money, as only the player with the highest dollar score will receive that value in the form of cash and be eligible to compete in the subsequent episode. The other two competitors receive consolation prizes based on their relative rank, with a prize of higher dollar value for second than third place (Trebeck and Barsocchini, 1990, p.57). As of May 16, 2002, the prior physical consolation prizes were changed to cash prizes with second place receiving US$2,000 and third place US$1,000 (Jeopardy!, 2015b).

Currently, the show ranks as the number two game show in syndication,[18] averaging 25 million viewers per week (Jeopardy!, 2015d). The production staff employs six researchers (cur-

[17] Before November 26, 2001, these values were US$100, 200, 300, 400, and 500. Similarly, the *Double Jeopardy!* clues included values of US$200, 400, 600, 800, and 1,000.

[18] Meaning the show is available to be licensed to television affiliates without ties to a specific network.

ACCEPTED MANUSCRIPT

rently two females and four males) and nine writers (currently two females and seven males) who are in charge of creating and assembling clues for the show (Jeopardy!, 2015c). The contestant selection process is initiated when interested individuals complete an online exam of 50 clues. Since 2006, online testing has become possible, which has contributed to expanding the contestant pool toward including more women, minorities, and students (Jeopardy!, 2015d).[19]

The show operates regular and tournament matches in four demographic categories: Kids (under the age of 13), teenagers (aged 13 – 17), college students (must be a full-time student and not have completed a bachelor's degree), and adults (over age 18). We access data from the regular show featuring adults (90.3 percent of episodes), but all our results are consistent when including the remaining categories.[20] According to Trebeck and Barsocchini (1990), approximately 250,000 people apply each year with 15,000 taking the first qualification exam, 1,500 qualifying for the show, and 500 being on air. Thus, only 0.2 percent of the initial applicants are eventually selected. Based on the examinations as part of the qualification process, contestants are "probably all equal in terms of knowledge" (Trebeck and Barsocchini, 1990, p.61) and a "random selection of material and contestants" is taking place (Trebeck and Barsocchini, 1990, p.40). Trebek states that ultimately it is a function of the categories that are randomly selected in a given episode that determines the winner.

ACCEPTED MANUSCRIPT

[19] The online qualification exam is available here: http://www.jeopardy.com/beacontestant/contestantsearches/practicetest/. Exams to be part of the show are offered periodically throughout the year and results are valid for 12 months (Jeopardy!, 2015a). Prior to 2006, paper examinations were given at various locations throughout the US.

[20] Results are not notably different when considering the remaining categories individually (kids, teenagers, or college students), but sample sizes become much smaller to conduct a meaningful regression analysis.

ACCEPTED MANUSCRIPT

Table AI: Summary statistics.

	Women		Men		p-value
	Mean	(Std. Dev.)	Mean	(Std. Dev.)	Women=Men
Panel A: All 4,268 episodes ($N = 12,804$, since 3 players/episode)					
N	5,127		7,677		
Winning episode	0.34	(0.47)	0.33	(0.47)	0.29
Panel B: Responding ($N = 749,433$)					
N	299,555		449,878		
Responding	0.33	(0.47)	0.33	(0.47)	0.75
STEM	0.06	(0.23)	0.06	(0.23)	0.42
$ value of clue	740	(508)	718	(501)	0.000***
Score	3,996	(4,424)	3,862	(4,240)	0.000***
Relative score	150	(837)	-100	(811)	0.000***
(2× score - score opp. 1 - score opp. 2)					
Panel C: Responding correctly ($N = 248,052$)					
N	99,211		148,841		
Responding correctly	0.85	(0.35)	0.85	(0.35)	0.46
STEM	0.06	(0.23)	0.06	(0.23)	0.56
$ value of clue	713	(492)	693	(486)	0.000***
Score	4,299	(4,712)	4,110	(4,457)	0.000***
Relative score	1,270	(8,597)	896	(8,205)	0.000***
(2× score - score opp. 1 - score opp. 2)					
Panel D: Daily Double clues ($N = 12,615$)					
N	5,010		7,605		
Wager as % of maximum	0.45	(0.30)	0.48	(0.31)	0.000***
STEM	0.06	(0.24)	0.07	(0.26)	0.12
Initial $ value of clue	1,045	(497)	1,013	(492)	0.000***
Score	6,120	(5,195)	5,770	(4,907)	0.000***
Relative score	3,291	(9,637)	2,834	(9,123)	0.007***
(2× score - score opp. 1 - score opp. 2)					
*DD*s before	2.67	(4.29)	4.95	(12.12)	0.000***

ACCEPTED MANUSCRIPT

Table AII: Robustness checks for estimating the likelihood to win a *Jeopardy!* episode.

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:		Females			Males	
	Logit[a]	Only 1^{st} show	No gaps[b]	Logit[a]	Only 1^{st} show	No gaps[b]
Dependent variable: *Winning episode (0/1, mean = 0.33)*						
# of male opponents	0.0341***	0.0263*	0.0305**			
	(0.0118)	(0.0139)	(0.0150)			
# of female opponents				0.0046	0.0081	-0.0050
				(0.0100)	(0.0128)	(0.0136)
Control variables[c]	yes	yes	yes	yes	yes	yes
Player-fixed effects			yes			yes
# of players	3,718	3,717	2,145	4,432	4,426	2,325
# of episodes	3,807	3,360	2,158	4,187	3,620	2,302
N	5,127	3,717	2,998	7,677	4,426	4,091

Notes: Standard errors clustered on the player level are displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a]Displaying marginal effects. [b]Only including episodes without any gaps (#4,451 to #7,084, since January 5, 2004). [c]Includes the number of previous *J!* episodes, in addition to year- and month-fixed effects.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table AIII: Robustness checks for estimating the likelihood to respond to a clue.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	Females				Males			
	Logit[a]	Only 1^{st} show	No gaps[b]	Alternative relative score[c]	Logit[a]	Only 1^{st} show	No gaps[b]	Alternative relative score[c]
Dependent variable: *Responding (0/1, mean = 0.331)*								
# of male opponents	0.0074***	0.0064***	0.0138**	0.0111**				
	(0.0020)	(0.0024)	(0.0055)	(0.0048)				
# of female opponents					-0.0002	0.0003	-0.0036	-0.0031
					(0.0017)	(0.0021)	(0.0043)	(0.0034)
Control variables[d]	yes	yes	yes	yes	yes	yes	yes	yes
Player-fixed effects			yes	yes			yes	yes
# of players	3,718	3,726	2,145	3,718	4,432	4,443	2,325	3,718
# of episodes	3,807	3,368	2,158	3,807	4,187	3,631	2,303	3,807
N	299,555	216,855	175,919	299,555	449,878	258,363	240,706	449,878

Notes: Standard errors clustered on the player level are displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a]Displaying marginal effects. [b]Only including episodes without any gaps (#4,451 to #7,084, since January 5, 2004). [c]Columns (4) and (8) employ an alternative measure for relative scores with player i's score minus the best opponent's score. [d]Includes binary indicators for STEM clues and the 20 most common categories, the $ value of the clue, the account balance of the contestant (both individual and relative to their opponents), the clue number in the current episode, a binary indicator for the *DJ!* round, an indicator whether the respondent has answered a clue incorrectly before, the number of previous *J!* episodes, and year- and month-fixed effects.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table AIV: Robustness checks for estimating the likelihood to respond correctly to a clue.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	Females				Males			
	Logit[a]	Only 1^{st} show	No gaps[b]	Alternative relative score[c]	Logit[a]	Only 1^{st} show	No gaps[b]	Alternative relative score[c]
Dependent variable: *Responding correctly (0/1, mean = 0.837)*								
# of male opponents	0.0037*	0.0020	0.0045	0.0040				
	(0.0019)	(0.0023)	(0.0049)	(0.0042)				
# of female opponents					-0.0035**	-0.0032	-0.0005	0.0015
					(0.0017)	(0.0020)	(0.0041)	(0.0032)
Control variables[d]	yes	yes	yes	yes	yes	yes	yes	yes
Player-fixed effects			yes	yes			yes	yes
# of players	3,718	3,726	508	782	4,432	4,443	839	1,369
# of episodes	3,807	3,368	1,157	1,899	4,187	3,631	1,900	3,359
N	99,211	72,645	26,419	42,439	148,841	90,106	48,518	86,673

Notes: Standard errors clustered on the player level are displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a]Displaying marginal effects. [b]Only including episodes without any gaps (#4,451 to #7,084, since January 5, 2004). [c]Columns (4) and (8) employ an alternative measure for relative scores with player i's score minus the best opponent's score. [d]Includes binary indicators for STEM clues and the 20 most common categories, the $ value of the clue, the account balance of the contestant (both individual and relative to their opponents), the clue number in the current episode, a binary indicator for the *DJ!* round, an indicator whether the respondent has answered a clue incorrectly before, the number of previous *J!* episodes, and year- and month-fixed effects.

ACCEPTED MANUSCRIPT

Table AV: Robustness checks for estimating the share of the maximum possible wager in *Daily Double* clues.

Sample:	(1) Females	(2)	(3)	(4)	(5) Males	(6)	(7)	(8)
	Alternative outcome variable[a]	Only 1st show	No gaps[b]	Alternative relative score[c]	Alternative outcome variable[a]	Only 1st show	No gaps[b]	Alternative relative score[c]
Dependent variable: *Responding correctly (0/1, mean = 0.837)*								
# of male opponents	183.8848***	0.0270***	0.0320**	0.0464***				
	(70.4345)	(0.0072)	(0.0159)	(0.0134)				
# of female opponents					-65.5245	-0.0161***	-0.0348***	-0.0309***
					(61.2901)	(0.0062)	(0.0130)	(0.0099)
Control variables[d]	yes	yes	yes	yes	yes	yes	yes	yes
Player-fixed effects			yes	yes			yes	yes
# of players	744	2,500	474	744	1,297	3,089	776	1,297
# of episodes	1,392	2,355	861	1,392	2,513	2,776	1,415	2,513
N	2,167	3,655	1,355	2,166	4,447	4,638	2,505	4,447

Notes: Standard errors clustered on the player level are displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a]Estimating the total wager, instead of the share of the maximum possible wager. [b]Only including episodes without any gaps (#4,451 to #7,084, since January 5, 2004). [c]Columns (4) and (8) employ an alternative measure for relative scores with player i's score minus the best opponent's score. [d]Includes binary indicators for STEM clues and the 20 most common categories, the $ value of the clue, the account balance of the contestant (both individual and relative to their opponents), the clue number in the current episode, a binary indicator for the *DJ!* round, an indicator whether the respondent has answered a clue incorrectly before, the number of previous *J!* episodes, and year- and month-fixed effects.