DISCUSSION PAPER SERIES

IZA DP No. 9669

Gender in *Jeopardy!*: The Role of Opponent Gender in High-Stakes Competition

Michael Jetter
Jay K. Walker

January 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Gender in *Jeopardy!*: The Role of Opponent Gender in High-Stakes Competition

Michael Jetter
University of Western Australia and IZA

Jay K. Walker
Niagara University

Discussion Paper No. 9669
January 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9669
January 2016

ABSTRACT

Gender in *Jeopardy!*: The Role of Opponent Gender in High-Stakes Competition[*]

Using 4,279 episodes of the popular US game show *Jeopardy!*, we analyze whether the opponents' gender is able to explain the gender gap in competitive behavior. Our findings indicate that gender differences disappear when women compete against men. This result is surprising, but emerges with remarkable consistency for the probability to (*i*) respond, (*ii*) respond correctly, and (*iii*) respond correctly in high-stakes situations. Even risk preferences in wagering decisions, where gender differences are especially pronounced, do not differ across gender once a woman competes against males. Using a fixed-effects framework, and therefore exploiting within-player variation only, confirms these findings. These results, derived from a large real-life setting, suggest that gender differences in performance and risk attitudes are not gender-inherent, but rather emerge in distinct social environments.

JEL Classification: D03, J10, J16

Keywords: competition, gender gap, risk preferences

Corresponding author:

Michael Jetter
Business School
University of Western Australia
35 Stirling Highway
Crawley, WA 6009
Australia
E-mail: mjetter7@gmail.com

[*] We are grateful to James Key and seminar participants at SUNY Buffalo for insightful discussions. All remaining errors are our own.

1 Introduction

Numerous studies find that, on average, women under-perform (e.g., Niederle, 2014) and exhibit more risk aversion in competitive situations (e.g., Byrnes et al., 1999). Persistent social phenomena, such as the gender wage gap (Goldin, 2014) or excessive risk-taking (Eckel and Füllbrunn, 2015), have been explained by such gender differences. But are these female-male disparities gender-inherent or do they only arise under specific circumstances? One hypothesis suggests that the gender of opponents can influence women's performance and risk attitudes (Booth et al., 2014). In the following pages, we analyze 8,169 contestants of the popular US game show *Jeopardy!* to test whether the opponents' gender indeed affects women's behavior in highly competitive situations. Contrary to previous studies with much smaller sample sizes, our results suggest that gender differences disappear once a woman competes against *males* only.

The idea that the gender of one's opponents affects behavior in general – and performance in particular – carries vital implications in a number of policy-relevant settings. For example, the debate over single- versus mixed-gender classes remains a persistent topic in the literature surrounding educational structures (Jackson, 2012; Park et al., 2013). Recently, this argument has again been the subject of heavy discussion in the UK.[1] Similarly, numerous work environments tend to be dominated by one gender, such as the finance industry or executive positions.[2] Especially in the finance industry, which remains a male-dominated domain, a distinctive degree of risk-taking may have contributed to the formation of the global financial crisis in 2007/08 (Eckel and Füllbrunn, 2015).

Unfortunately, analyzing potential effects of the opponents' gender on women's performance in competitive situations has been challenging. In real life, workers are likely to self-select into a desired gender environment, which can introduce a non-trivial endogene-

[1]For example, see articles in *The Telegraph* (Boult, 2016; Turner and de Normann, 2016), *The Guardian* (Lott-Lavigna, 2016), or the *Cambridge News* (Cambridge News, 2016) for different opinions.

[2]See Eckel and Füllbrunn (2015) for the share of women in finance; Bertrand and Hallock (2001) point out that only 2.5 percent of executive positions in the United States have been filled with women. Similarly, the share of females in IT- and math-related occupations remains low, sometimes less than 20 percent, as highlighted by the US Department of Labor (DoL, 2016).

ity bias when regressing performance indicators on variables that capture the opponents' gender. Further, individual outcomes are difficult to observe and isolate in most occupations. As a response, researchers have largely turned to experimental studies. Although causality can be isolated more effectively in laboratory experiments, these studies usually rely on fewer observations and the generalizability of results remains a concern (e.g. Levitt and List, 2007; Niederle, 2014; Alesina and Giuliano, 2015).

We analyze a real-life situation where people are assigned male or female opponents, using information from the performance of 8,169 *Jeopardy!* contestants in 749,433 observations. Contestants have no control over whether they compete against *a*) two females, *b*) one female and one male, or *c*) two males. We evaluate whether the gender of the opposition affects women's probabilities to (*i*) respond, (*ii*) respond correctly, and (*iii*) respond correctly in particularly high-stakes *Daily Double* clues. Further, we test whether the gender of women's opponents influences wagering decisions in *Daily Double* opportunities. As prize money averages US$20,648 (in 2015 US$) per episode, incentives are substantial and players are likely to reveal their true competitive behavior.[3] In additional estimations, we also analyzed the effect of gender on men's performance, but the corresponding results are generally not statistically significant and no clear pattern emerges.

To relate our setup to the closest findings in the literature, Table 1 summarizes some of the most recent works discussing gender differences in competitive behavior and risk preferences. By no means do we claim this list to be complete, but it illustrates some notable insights. Most importantly, the majority of studies have been experimental with sample sizes ranging from less than 100 to about 500 and payoff structures markedly lower than in *Jeopardy!*.

Beginning with competitive situations, the corresponding results establish the notion that women are, on average, under-performing in competition and less likely to select into competitive situations (Niederle and Vesterlund, 2007). Potential explanations relate to

[3]These monetary incentives compare well to an experiment conducted by Ariely et al. (2009) in India, who emphasize large stakes. The prize for the winner of an average *Jeopardy!* episode equals approximately 40 percent of GDP per capita in the US. The median payoff amounts to US$19,752, whereas the maximum amount a winner managed to take home is US$93,993.

Table 1: Most recent articles on the effect of gender and gender composition on performance and risk-taking (listed chronologically).

Authors (Year)	Empirical Design	Sample	N	Payoffs	Main Findings
Panel A: Gender differences in competitiveness					
Jetter and Walker (2015)	Real life	Tennis players	107,566	Up to US$1 million	Female tennis players are equally likely to perform well in high stakes situations than male players
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
Lavy (2013)	Real life	Teachers	724	Maximum possible payoff US$7,500	No gender differences in competitiveness over cash bonus for class performance
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
Fryer and Levitt (2010)	Real life	School kids	9,481	Standardized testing environment	Substantial gender gap in math scores develops over school career
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)
Niederle and Vesterlund (2007)	Experiment	Adults	77	Average payoff $19.80	Women are less likely to select into competitive situations than men
Gneezy and Rustichini (2004)	Experiment	School kids	140	None	Competition improves performance for boys, not for girls
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 risk preferences					
Säve-Söderbergh and Lindquist (2016)	Real life	Adults & kids	556	Average $1,850 adults	Women wager less when competing against men
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	$15,000 Australian	Mixed evidence of greater risk aversion by women in Australian *Deal or No Deal* show

the task at hand (Shurchkov, 2012; Günther et al., 2010), the time given to learn (Cotton et al., 2013; Shurchkov, 2012), or team versus individual competition formats (Healy and Pate, 2011). Note that several non-experimental studies listed here find no remarkable gender differences in competitive behavior (Lavy, 2013; Jetter and Walker, 2015), contrary to most laboratory settings. Similarly, the results from our study reveal no statistically significant gender differences in competitive situations. However, we do find substantial differences when distinguishing by gender of the opposition: a woman performs equally to men *only* when competing against males. Otherwise, she under-performs.

Concerning risk preferences, competing against men has usually been found to discourage risk-taking by females (Booth and Nolen, 2012b; Booth et al., 2014), as summarized in Panel B of Table 1. Our findings stand in stark contrast to these results, indicating that a woman exhibits the same risk preferences as a man when competing against males. In the presence of other women, however, females play much more conservatively and wager significantly less in *Daily Double* clues. In terms of magnitude, our most complete estimation suggests women to wager 3.3 percentage points less than men when competing against other females. In contrast, she wagers 0.4 percentage points *more* than a man when competing against two males. It is important to highlight that our results remain remarkably consistent after the inclusion of a comprehensive set of control variables, such as her previous *Jeopardy!* performance or the nature and category of the clue.

Finally, incorporating player fixed effects, therefore exploiting within-player dynamics only, confirms our findings and, if anything, the corresponding quantitative interpretation becomes stronger. The same female contestant performs better and wagers significantly more when competing against two males, as opposed to playing against at least one female.

Our results carry implications for several streams of literature and policy areas. First, work environments dominated by females may produce under-performance in high-stakes competitive situations. Second, increasing the share of women in male-dominated occupations, such as the finance industry, may not necessarily alter the overall degree of risk-taking, as suggested by Eckel and Füllbrunn (2015). One woman in an all-male environment may simply assimilate to the males' behavior. By design, our study is limited to analyzing three contestants and further studies may be needed to evaluate at which gender-ratio of competitors behavioral differences emerge. Third, our findings contribute to the debate about nature versus nurture, implying that differences in competitive behavior are driven by surrounding characteristics (the gender of opponents), rather than being gender-inherent. This may favor the general concept of "social learning" over inherent gender differences in explaining gender differences in competitive situations and risk preferences, as described by Booth and Nolen (2012b).

The paper proceeds as follows. Section 2 provides an overview of the existing literature, in addition to a detailed description of *Jeopardy!*. Section 3 presents our sample data and describes our empirical methodology. In Section 4, we describe our findings. Section 5 places our findings in context to the existing literature. Finally, Section 6 concludes with policy implications of our findings.

2 Background

This section provides a brief summary of the literature surrounding gender differences in competitive behavior and risk preferences, focusing on the importance of the opponents' gender. Given the large extent of these literatures, we restrict ourselves to the most prominent and recent contributions. Finally, we discuss the structure and history of *Jeopardy!* to illustrate how this setting is particularly well suited to study the effect of gender composition on performance indicators.

2.1 Gender Differences in Competitive Behavior

In a seminal paper, Gneezy et al. (2003) show females to be less effective than males in competitive environments. Their results imply that raising the degree of competitiveness leads to an improvement in men's performance, but not in women's. In related studies, Gneezy and Rustichini (2004), Niederle and Vesterlund (2010), and Reuben et al. (2015), among others, find girls and women to under-perform in competitive situations. Croson and Gneezy (2009) provide an excellent summary of this literature. Overall, gender differences in competitive behavior have been observed in a variety of settings and at different ages.

Recently, a number of potential explanations have been proposed. For example, Shurchkov (2012) and Günther et al. (2010) suggest task stereotypes and time constraints as possible drivers. Women may under-perform in high-pressure situations and when time is limited. Related to that hypothesis, Cotton et al. (2013) find girls to close the performance gap to boys in a repeated game, as the gender gap only emerges in the first period.

A key question for policymakers then asks whether this heterogeneity is gender-inherent, acquired in society ("social learning"), or dependant on particular situations. Andersen et al. (2013) conduct a series of experiments to find that in both matrilineal and patriarchal societies girls are equally competitive as boys until puberty. However, in male-dominated cultures girls become less competitive thereafter. Booth and Nolen (2012b) use the term "social learning" in this context. In what follows, our results confirm the notion that performance differences may not be gender-inherent, but rather depend on social environments.

2.2 Gender Differences in Risk Preferences

Finally, an equally well-developed stream of literature has analyzed gender differences in risk preferences. Byrnes et al. (1999) summarize the corresponding literature, concluding males to generally be less risk averse, but this gap seems to narrow over time. Nevertheless, gender differences in risk preferences have consistently been shown in a variety of settings, such as decisions about smoking and seat belt usage (Hersch, 1996) or financial investment (Jianakoplos and Bernasek, 1998; Eckel and Füllbrunn, 2015).

Closely related to the design of our study, De Roos and Sarafidis (2010) analyze risk preferences by gender on the game show *Deal or No Deal*. Their results produce mixed evidence for the claim that females are more risk averse than males. Kelley and Lemke (2015) note women and men appear to weigh subjective probabilities differently when considering a gamble, using data from the game show *Cash Cab*. The closest predecessor to our paper comes from Säve-Söderbergh and Lindquist (2016) who study risk preferences in the Swedish edition of *Jeopardy!*. We will relate to their findings in detail in section 5. In general, several researchers have used data from game shows to analyze people's behavior (e.g., Gertner, 1993, Metrick, 1995, Levitt, 2004, Antonovics et al., 2005, or Post et al., 2008).

2.3 Gender of Opponents

Recently, the gender of one's competitors has been suggested as a potential driver of the described gender differences in performance and risk preferences. This hypothesis has mostly been tested in experimental studies, as it is difficult to find real-life competitive situations where *a*) the gender of one's opponent is randomly assigned, *b*) individual outcomes are clearly observable, and *c*) a sizeable sample is available. In reality, people are likely selecting into gender environments they are comfortable with.

In a series of seminal experimental studies, Alison Booth and Patrick Nolen find girls and female college students to take on higher risks in single-gender environments (Booth and Nolen, 2012a; Booth and Nolen, 2012b; Booth et al., 2014). Interestingly, Datta Gupta et al. (2013) find that once women can choose the gender of their competitors, their willingness to compete increases, but this does not fully explain the gender gap.[4]

Although these results from experimental studies are powerful in randomizing the gender of one's opposition, it is difficult to draw generalizable conclusions from these. External validity is one of the key critiques of experimental studies that usually have to rely on relatively small sample sizes, in addition to artificial and many times diminished

[4]Other works relating to the willingness to compete are provided by Shurchkov (2012), Gneezy and Pietrasz (2013), and Sutter and Glätzle-Rützler (2014), among others.

incentive structures. Antonovics et al. (2009) focus on the comparability of results obtained from laboratory settings with real-life situations. They conclude that findings are comparable when the offered stakes are high (above US$50) and when players are young (under the age of 33). Note that a minimum payoff of US$50 can quickly become expensive if a researcher would like a sizeable sample. Most experimental studies have found it difficult to cross that payoff threshold (see Table 1).

As an alternative, some studies have used data from game shows. In particular, Lindquist and Säve-Söderbergh (2011) and Säve-Söderbergh and Lindquist (2016) use data from the Swedish edition of *Jeopardy!* to analyze whether the opponents' gender affects women's risk-taking behavior. Consistent with the experimental evidence described above, they find women to take on less risk when competing against males. Overall, however, non-experimental studies to test the effect of the opponents' gender on performance and risk-taking have been scarce.

2.4 *Jeopardy!*

On September 10, 1984, *Jeopardy!* started its current run on television. Each episode hosts three contestants and a maximum of 61 clues. Contrary to conventional game shows, the *Jeopardy!* host, Alex Trebek, reads out 'clues' and whoever contestant pushes a buzzer first has to pose the correct question to the 'clue.' In turn, whoever provides the correct question to the clue is able to decide which clue will be selected next. Throughout the paper, we will use the terminology of 'responding' to clues to facilitate readability. The show consists of three rounds: *Jeopardy!*, *Double Jeopardy!*, and *Final Jeopardy!*.

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.[5] In each *Jeopardy!* round, one *Daily Double* clue is hidden, whereas two are included in the *Double Jeopardy!* round.

[5]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.

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 one 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,[6] averaging 25 million viewers per week (Jeopardy!, 2015d). The production staff employs six researchers (currently two female and four male) and nine writers (currently two female and seven male) 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 been possible, which has contributed to expanding the contestant pool toward including more women, minorities, and students (Jeopardy!, 2015d).[7]

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

[7] 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 geographic locations throughout the United States.

The show operates regular and tournament matches in four demographic categories: kids (under the age of 13), teens (aged 13 – 17), adults (over age 18), and college students (must be a full-time student and not have completed a bachelor's degree). 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. 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). Trebek states that ultimately it is a function of the categories that are randomly selected for a given show that determines a winner. The clues' topics are drawn from general knowledge categories and our analysis will control for the most common categories.

Our study is not the first to employ *Jeopardy!* data. Most notably, Metrick (1995) analyzes 1,150 *Final Jeopardy!* round decisions from 1989 to 1992, estimating risk preferences. However, his focus is not on gender differences or gender composition of opponents. He et al. (2008) use *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, too, do not focus on the gender of opponents. Finally, Lindquist and Säve-Söderbergh (2011) and Säve-Söderbergh and Lindquist (2016) use data from the Swedish edition of *Jeopardy!* to analyze the effect of opponents' gender on girls' and women's risk preferences in *Daily Double* clues.

3 Data and Methodology

3.1 Data Description

As of June 5, 2015, the *J! Archive* website, a fan-created archive of *Jeopardy!* episodes, contains full information for 4,279 complete episodes, including 8,169 players and 254,079 clues. Employing a data specialist, we extracted all information of all available episodes. While the dataset is extensive, it does not include all *Jeopardy!* episodes. In fact, recent seasons contain information on every episode (since January 5, 2004), but earlier seasons often have missing episodes on the website. However, if an episode is present, all corresponding information is available. After checking the episode numbers and available data, we do not find evidence of systematic omissions from the archive. Nevertheless, limiting the sample to the episodes #4,451 to #7,084 (since January 5, 2004), where the data exhibits no gaps, produces consistent findings. The corresponding results are available in Tables AII to AIV.

For every clue given, the website contains information about the three contestants' full names, their accumulated prize money, the category of the clue, and the sequence of clues. 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., Frank and Adam are male; Alison and Alyssa are female) 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 *Jeopardy!* contestant. The same follows if names are abbreviated (e.g., A.C. Hawley). This strategy allows us to allocate a binary indicator for female contestants to all of the 8,169 contestants.

3.2 Four Scenarios of Study

Throughout the analysis, we focus on four distinct competitive scenarios occurring with the *Jeopardy!* and *Double Jeopardy!* rounds.[8] First, we consider each contestant's probability to answer a clue. Thus, for each clue we obtain three data points (one for each contestant), producing a total sample size of 749,433 observations. Second, once a candidate taps the buzzer we evaluate the probability of responding correctly. 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). This second scenario produces a sample of 248,052 observations.

Third, we analyze performances in particularly high-stakes situations: responding to *Daily Double* clues. In these clues, stakes (as wagered by the respective contestant) average US$1,922, whereas the average stakes in a regular clue come out to be US$711. Overall, this produces a sample of 12,615 observations. Fourth and final, we analyze the wagering decision in these *Daily Double* scenarios. Specifically, we divide the chosen wager by the maximum possible wager to estimate what percentage of the allowed sum is wagered (akin to Säve-Söderbergh and Lindquist, 2016). Everything else equal, a higher percentage indicates higher stakes and therefore higher risks.

Table 2 displays the dependent variables of all four scenarios by gender. At first glance, we observe no overall gender differences in the probabilities to respond or to respond correctly. However, women choose to wager less than men in *Daily Double* situations (45.2 percent of their available maximum, compared to 47.5 percent). These basic statistics are consistent with benchmark results from numerous studies, suggesting women to be more risk-averse than men (e.g., Eckel and Füllbrunn, 2015).

[8] We exclude *Final Jeopardy!* rounds, as all contestants naturally respond to that clue. In addition, wagering decisions in the *Final Jeopardy!* round can follow a number of different strategies, depending on the account balances of all contestants (see Metrick, 1995).

Table 2: Summary statistics by gender.

Variable	(1) Females Mean (N)	(2) Males Mean (N)	(1) = (2) p-value
Answering	0.331 (299,555)	0.331 (449,878)	0.754
Answering correctly	0.855 (99,211)	0.853 (148,841)	0.457
Answering correctly to *Daily Double* clue	0.637 (5,010)	0.649 (7,605)	0.174
Wager in Daily Double as % of maximum	0.452 (5,010)	0.475 (7,605)	0.000***

3.3 Descriptive Statistics of Gender Composition

Figure 1 focuses on the females in all four settings and the gender composition of their opponents. A woman rarely competes against two other women, whereas in about half of the observations she competes against two males. Throughout our analysis, we will consider the mixed opposition (one female, one male opponent) as the reference point. Nevertheless, whichever category forms the reference group, our results are consistent in highlighting the importance of a woman competing against two males.

Table 3 displays the four respective performance indicators when women compete against two males, compared to any other gender composition (female/male and female/female). Women are performing better in all three performance measures and are taking higher risks when competing against males only. Nevertheless, these basic descriptive statistics should of course be considered with care. It is possible that confounding factors are driving these differences, such as the category of the clue, previous performance, current standings (individual and relative to opponents), or time-specific aspects.

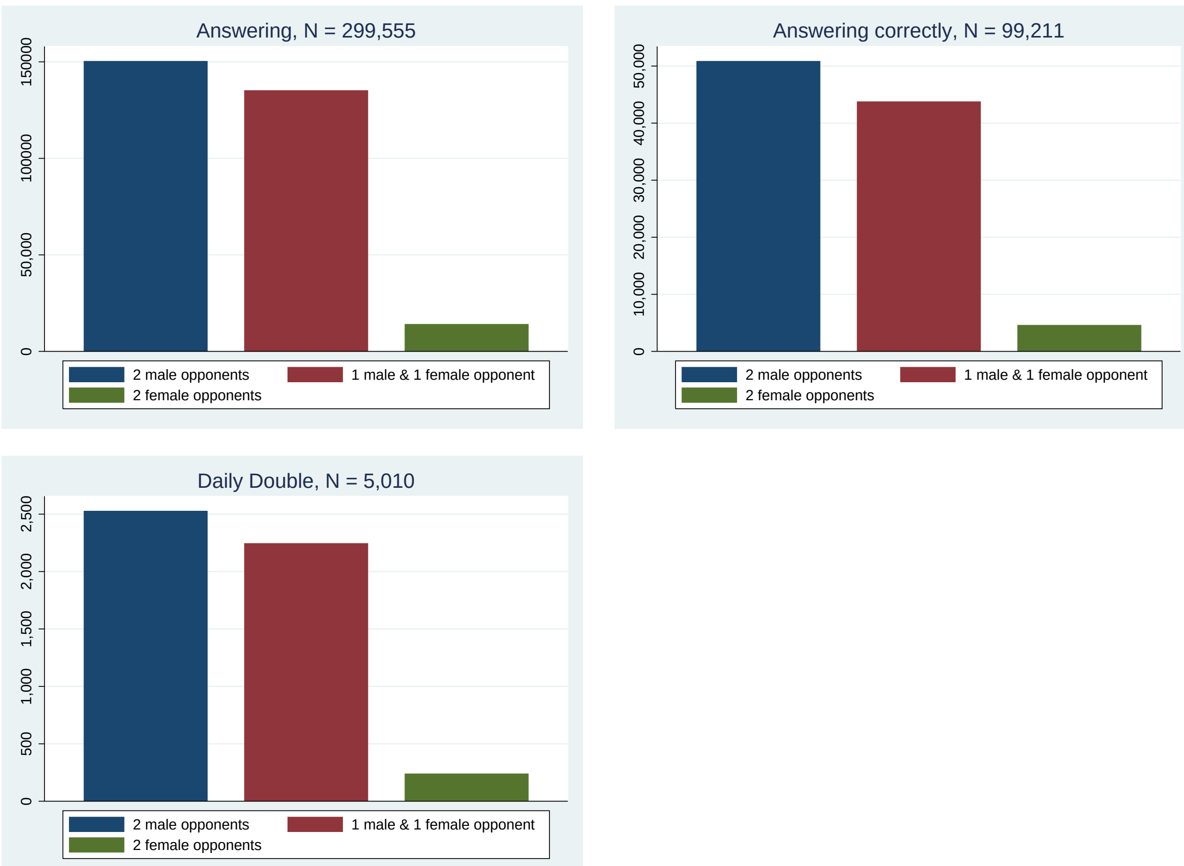

Figure 1: Gender composition of the opposition for all females in the sample.

Table 3: Summary statistics of females by gender of the opposition.

Variable	(1) 2 male opponents Mean (N)	(2) 1 or 2 female opponents Mean (N)	(1) = (2) p-value
Answering	0.338 (150,322)	0.324 (149,233)	0.000***
Answering correctly	0.857 (50,850)	0.852 (48,361)	0.016**
Answering correctly to *Daily Double* clue	0.658 (2,527)	0.616 (2,483)	0.002***
Wager in Daily Double as % of maximum	0.465 (2,527)	0.440 (2,483)	0.004***

3.4 Control Variables

Given the information provided on the *J! Archive* website, we control for a number of potentially confounding factors in all four scenarios. In particular, we incorporate binary indicators for STEM clues (science, technology, engineering, and mathematics), 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 round, a binary indicator for the *Double Jeopardy!* round, an indicator whether the respondent has answered a clue incorrectly before, and year fixed effects.[9] In addition, we incorporate player fixed effects, therefore relying on within-player variation only. We now briefly discuss the relevance of each factor in estimating our four performance indicators.

[9]We browsed all categories, manually sorting categories 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.

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. In general, it is possible that some categories are generally preferred by either gender and our analysis includes binary indicators for the 20 most common categories. Second, a line of research focuses on the size of an expected payoff, which relates 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. Thus, including the stakes of the clue may constitute a valuable performance indicator. In addition, the *Jeopardy!* outline proposes that the degree of difficulty increases with the associated dollar value (Provencher, 2011).

Third, although *Jeopardy!* 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. Thus, we incorporate the current account balance of the respective participant into the regression analysis. This variable may not only reflect the contestant's capabilities with respect to *Jeopardy!*, but also their degree of confidence in the ongoing show. In addition, to capture the relative standing of a player, we incorporate a variable relating one's current account balance to their opponents'. This caters to the notion that prior performance of competitors may influence behavior in competitive tasks (Smith, 2013, produces evidence from spelling bee contests). To retain all possible observations, we choose an additive formula: $2 \times \textit{own balance} - \textit{balance}_1 - \textit{balance}_2$, where subscripts denote opponents. 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 other relative performance indicators.

Fourth, we incorporate a variable capturing whether a contestant has responded incorrectly to a clue before, relating to the same concept. All results are robust when including a binary indicator for whether a contestant has responded correctly before, following Post et al. (2008) and He et al. (2008). The corresponding results are referred to Tables AII to AIV.

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 *Double Jeopardy!* 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 coefficients on gender-related regressors may be biased if such a variable is not captured.

Sixth, we include year fixed effects. Figure 2 visualizes the share of female contestants over time, as well as the fraction of shows that feature one female against two males. Note that, although the participation of women has fluctuated from around 22 percent to almost 50 percent, we observe more women in the show since 2006, when online testing has been introduced (see section 2.4 for details). As for gender composition of the show, the 1-female-2-males combination has been especially prominent in the first year of the show and has since fluctuated between 40 percent and over 70 percent. A number of factors could have driven these developments over time and we wish to ensure that no time-specific unobservables are influencing our analysis. Note that year fixed effects also control for the mentioned doubling of prize money at the end of 2001.

Finally, in an extension discussed in section 4.5 we further incorporate player fixed effects, thereby relying on within-player variation only. Detailed summary statistics of all control variables are provided in Table AI.

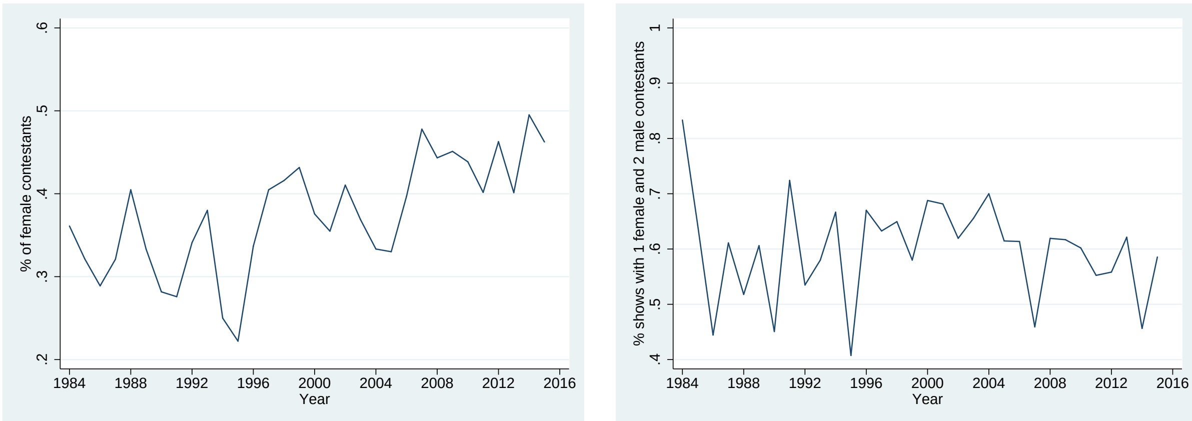

Figure 2: Share of female contestants over time (left) and share of shows with one female contestant (right).

3.5 Methodology

Our econometric strategy follows standard regression methodologies in the associated literature. To analyze probabilities of responding and responding correctly, we first apply a conventional logit regression framework, clustering errors on the player level. Throughout our analysis, we will display marginal effects of all associated coefficients to facilitate interpretation. Nevertheless, all our results are robust when applying a probit framework and are referred to Tables AII to AIV. To incorporate unobservable characteristics of each individual player, such as individual ability, we then move to a fixed effects framework.

Our analysis of wagering decisions applies an OLS framework, again clustering errors on the player level. Throughout all our estimations, we focus on the coefficients of interest, although detailed results of the control variables are displayed in Tables AII to AIV.

4 Empirical Findings

Tables 4 through 7 display regression results for all four performance indicators. In each Table, columns (1) and (2) analyze the full sample, checking for basic gender differences. Columns (3) and (4) focus on the female subsample, distinguishing between the gender composition of their opponents. Further, we form subsamples including all males and distinguishing women by the gender of their opponents. Column (7) of each Table concludes with an estimation of the full sample, including a binary term for females who happen to compete against two males. Finally, Tables 8 and 9 present our results from incorporating player fixed effects.

4.1 Answering

We begin by analyzing the probability to respond to a clue. Note that, in the context of competitive behavior, responding could reflect the pure choice to respond, but it may also involve the quickness with which a contestant hits the buzzer following the prompt. Note that the prompt to respond is made available only after the full clue has been read (Trebeck and Barsocchini, 1990, p. 58-60). Thus, even though a contestant may want to respond, they may not get to do so if an opponent is faster to tap the buzzer after responses are allowed. Nevertheless, both explanations (speed and knowledge) can arguably be considered as behavior in a competitive situation.

Table 4 displays the corresponding results from logit regressions, estimating the probability that a contestant responds to a clue. Consistent with the raw descriptive statistics displayed in Table 2, we observe no gender differences in the probability to respond, even after including all control variables. Women are as likely as men to respond to a given *Jeopardy!* clue.

Columns (3) and (4) focus on the females in our sample, testing whether the gender composition of their contestants matters in their propensity to respond. Indeed, substantial heterogeneity emerges. If women compete against men only, they are more likely to respond, compared to the scenario where at least one other woman forms part of the group. This result is statistically significant on the one percent level. Column (4) further introduces a binary indicator for a single-sex environment (only females), but the result related to an all-male composition of opponents remains unchanged. Thus, *the absence of women* increases the likelihood of a woman to respond. In terms of magnitude,

Table 4: Dependent variable is whether contestant responds. Displaying marginal effects from logit regressions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full Sample		Females only		Subsamples[b]		Full Sample
Female	0.000 (0.002)	-0.000 (0.002)			0.004* (0.002)	-0.005*** (0.002)	-0.005** (0.002)
Control variables[a]		yes	yes	yes	yes	yes	yes
2 male opponents			0.010*** (0.002)	0.010*** (0.002)			
2 female opponents				0.001 (0.005)			
Female × 2 male opponents							0.009*** (0.002)
# of players	8,169	8,169	3,726	3,726	6,731	6,326	8,169
% female players	39.97	39.97	100	100	25.04	24.91	39.97
N	749,433	749,433	299,555	299,555	600,200	599,111	749,433
Log lik.	-475,797	-471,953	-188,446	-188,446	-378,544	-376,791	-471,938

Notes: Standard errors clustered on the player level 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 round, a binary indicator for the *Double Jeopardy!* round, an indicator whether the respondent has answered a clue incorrectly before, and year fixed effects. [b]Column (5) includes all males and those females who compete against 2 males. Column (6) includes all males and those females who compete against 1 or 2 females.

women are one percentage point more likely to respond when competing against males only, compared to competing against at least one other woman.

Moving to columns (5) and (6), we consider two distinct subsamples to analyze how much the gender of a woman's opponents influences her performance. Specifically, we exploit the random assignment of all contestants into competition with people from either gender. Column (5) forms a subsample including all males and those females who happen to compete against two males. Indeed, the results show a woman's likelihood of responding is actually *higher* than a man's, as soon as she is competing against two males.

Column (6) considers a subsample of males and only those females who happen to compete against at least one other woman. In this case, a woman is 0.5 percentage points *less* likely to respond than an average man and the gender gap emerges with force. This result is statistically significant on the one percent level.

Finally, column (7) returns to using the full sample, introducing a term that takes on the value of one if a female competes against two males. The pure gender coefficient is now negative, statistically significant on the one percent level, and indicates women to be five percentage points less likely to respond. However, when a woman competes in an otherwise all-male environment, the net effect turns to +0.004 (-0.005 plus 0.009) and the overall statistical effect is undistinguishable from zero.

In sum, Table 4 produces clear differences in a woman's probability to respond to a *Jeopardy!* clue, depending on the gender of her opposition. A woman's performance, measured as the probability to respond, improves significantly when competing against two males.

4.2 Answering Correctly

After the probability of responding, we now evaluate the performance of *Jeopardy!* contestants in answering clues correctly. Thus, this sample considers those 248,052 observations where the respective contestant chose to respond, with the corresponding results displayed in Table 5. As in the first scenario, we find no basic gender differences in the probability to respond correctly in columns (1) and (2), even after the inclusion of the comprehensive set of control variables. However, women once again fare better when competing against males only. In this case, women are 0.4 percentage points more likely to answer correctly – a result that is statistically relevant on the ten percent level.

Table 5: Dependent variable is whether contestant responds correctly. Displaying marginal effects from logit regressions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full Sample		Females only		Subsamples[b]		Full Sample
Female	0.001 (0.002)	-0.000 (0.001)			0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)
Control variables[a]		yes	yes	yes	yes	yes	yes
2 male opponents			0.004* (0.002)	0.004* (0.002)			
2 female opponents				-0.002 (0.005)			
Female × 2 male opponents							0.004* (0.002)
# of players	8,169	8,169	3,726	3,726	6,731	6,326	8,169
% female players	39.97	39.97	100	100	25.04	24.91	39.97
N	248,052	248,052	99,211	99,211	199,691	197,202	248,052
Log lik.	-103,149	-100,662	-40,180	-40,180	-80,859	-80,250	-100,660

Notes: Standard errors clustered on the player level 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 round, a binary indicator for the *Double Jeopardy!* round, an indicator whether the respondent has answered a clue incorrectly before, and year fixed effects. [b]Column (5) includes all males and those females who compete against 2 males. Column (6) includes all males and those females who compete against 1 or 2 females.

Columns (5) and (6) consider the familiar subsamples, first analyzing a sample of men and those women with male-only competition. Here again, we find no statistically relevant difference between males and females – if anything, women are more likely to respond correctly. However, a woman's probability of responding correctly decreases as soon as other women are competing, although the corresponding coefficient is not statistically significant on conventional levels. Overall, the same trend from Table 4 is maintained.

Finally, column (7) returns to the full sample and, similar to the results from Table 4, highlights that gender composition matters for women. Generally, a woman is 0.2 percentage points less likely to respond correctly, although this effect is not significant on conventional levels of statistical relevance. In turn, her performance improves when competing against men only.

4.3 Answering Correctly in Daily Double Clues

The third setting aimed at analyzing performance focuses on *Daily Double* clues. We choose these situations because the expected payoffs are substantially elevated, compared to the average *Jeopardy!* clue. Throughout our sample, the wagered value of *Daily Double* clues averages US$1,922 as opposed to US$711 in a regular clue.[10]

Table 6 displays results from estimating the probability to respond correctly to these high-stakes clues. In the full sample, females are marginally less likely to respond correctly, a result that is consistent with previous findings in the associated literature (Gneezy et al., 2003; Gneezy and Rustichini, 2004; Niederle and Vesterlund, 2010). Nevertheless, the coefficient is not significant on conventional levels of statistical relevance.

Gender composition continues to play a decisive role, as women are three percentage points more likely to respond correctly when paired against males (columns 3 and 4). Further, column (5) shows that the gender performance gap in high-stakes situations disappears if a woman is randomly assigned to compete against men only. In fact, the coefficient on *female* turns positive. However, when comparing men's performance to those women who happen to compete against at least one other female, the gender gap emerges with force: women are 2.7 percentage points less likely to respond correctly. Once again, the presence of other women seems to notably diminish the performance of a woman.

Column (7) reiterates these findings in the full sample. In general, women are 2.8 percentage points less likely to respond correctly in high-stakes clues, compared to men.

[10] The median value of *Daily Double* clues is US$1,500, as opposed to US$600 for regular clues.

Table 6: Dependent variable is whether contestant responds correctly to *Daily Double* clue. Displaying marginal effects from logit regressions.

	(1) Full Sample	(2) Full Sample	(3) Females only	(4) Females only	(5) Subsamples[b]	(6) Subsamples[b]	(7) Full Sample
Female	-0.012 (0.009)	-0.012 (0.009)			0.005 (0.011)	-0.027** (0.011)	-0.028*** (0.011)
Control variables[a]	yes	yes	yes	yes	yes	yes	
2 male opponents			0.032** (0.014)	0.030** (0.014)			
2 female opponents				-0.013 (0.033)			
Female × 2 male opponents							0.033** (0.013)
# of players	6,075	6,075	2,683	2,683	4,967	4,780	6,075
% female players	44.16	44.16	100	100	31.69	29.02	44.16
N	12,615	12,615	5,010	5,010	10,132	10,088	12,615
Log lik.	-8,209	-8,067	-3,189	-3,189	-6,435	-6,468	-8,064

Notes: Standard errors clustered on the player level 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 round, a binary indicator for the *Double Jeopardy!* round, an indicator whether the respondent has answered a clue incorrectly before, and year fixed effects. [b]Column (5) includes all males and those females who compete against 2 males. Column (6) includes all males and those females who compete against 1 or 2 females.

However, that difference disappears once a woman competes in an otherwise all-male environment.

4.4 Wagering in Daily Double Clues

Our final setting studies the wagering behavior of contestants who happen to choose a *Daily Double* clue. Based on a number of previous findings (e.g., Byrnes et al., 1999; Eckel and Füllbrunn, 2015), we would expect women to be more risk-averse than men in their wagering decisions. Indeed, columns (1) and (2) of Table 6 firmly support this notion. On average, a woman wagers 1.5 percentage points less of her available maximum betting amount than men, once the familiar control variables are accounted for.

Table 7: Results from OLS regressions estimating the share of maximum wager in *Daily Double* clue.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full Sample		Females only		Subsamples[b]		Full Sample
Female	-0.022*** (0.006)	-0.015*** (0.004)			0.004 (0.005)	-0.033*** (0.005)	-0.033*** (0.005)
Control variables[a]	yes	yes	yes	yes	yes	yes	
2 male opponents			0.036*** (0.007)	0.035*** (0.007)			
2 female opponents				-0.009 (0.015)			
Female × 2 male opponents							0.037*** (0.006)
# of players	6,075	6,075	2,682	2,682	4,967	4,779	6,075
% female players	44.16	44.16	100	100	31.69	29.02	44.16
N	12,615	12,615	5,010	5,010	10,132	10,088	12,615

Notes: Standard errors clustered on the player level 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 round, a binary indicator for the *Double Jeopardy!* round, an indicator whether the respondent has answered a clue incorrectly before, and year fixed effects. [b]Column (5) includes all males and those females who compete against 2 males. Column (6) includes all males and those females who compete against 1 or 2 females.

Once we consider differences among women, depending on the gender of their competitors, we find the same heterogeneity as in the previous settings. Women are wagering considerably more in an otherwise all-male environment. Further, columns (5) and (6) reveal how important the role of gender composition really is: if a woman competes against men only, her risk-taking behavior is not in any way distinguishable from that of men. In fact, the respective coefficient even turns positive, hinting at women potentially taking higher risks than men when competing against an all-male field. Nevertheless, the corresponding coefficient remains far from conventional levels of statistical importance. Contrary to that, column (6) paints a different picture. In a sample of males and only those females who happen to compete against at least one other female, the gender gap in risk preferences emerges with force. In this case, the gap increases to over twice the initial size in column (2) and women wager 3.3 percentage points less than men, on average.

Finally, column (7) displays results from analyzing the full sample. Consistent with the previous findings, no gender gap emerges if a woman competes against two males, as the net effect of being female in that situation reaches a value of +0.004. If anything, the results in Table 6 are stronger in qualitative and quantitative terms than those for the previous scenarios of responding and responding correctly. Thus, gender composition of a woman's competitors may substantially influence her risk-taking behavior.

4.5 Including Player Fixed Effects

Up to now, our analysis relies on variation within and across contestants in a pooled setting. However, although we incorporate a number of potentially confounding factors and errors are clustered on the player level, individual particularities could drive our findings. For instance, *Jeopardy!* producers may be able to identify promising candidates beforehand and then choose promising female candidates to compete against two males. None of our readings about the show would suggest this, but it is also likely that producers would not openly admit to such a strategy. In general, unobservable characteristics on the individual level may differ between those women who compete against two males and the remaining female candidates.

To alleviate such concerns, we now move to a fixed effects framework. These regressions only exploit within-player variation, therefore controlling for any individual heterogeneity. In other words, the derived coefficient for women with two male opponents only reflects information from those females who have competed against two males *and* against at least one female in another episode. By design, such 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.

Thus, for our first three settings of binary outcome variables, we first replicate the most complete estimation (column 7 of each respective Table) in an OLS model to see whether results are comparable. Then, we employ the fixed effects framework. (Note that the pure gender variable becomes obsolete in a fixed effects framework.) Table 8 displays the corresponding findings from analyzing the probabilities to answer (columns 1 to 3) and to answer correctly (columns 4 to 6). The first respective column simply replicates the most complete logit regression, before moving to a pooled OLS framework and then to including fixed effects.

First, notice that the OLS analyses in columns (2) and (5) closely replicate the results from logit regressions. Incorporating fixed effects then produces telling results, as the derived coefficients for the variable of interest are confirmed in their statistical importance. In fact, the respective magnitudes are even increased. For the decision to respond, the regression produces a coefficient of 0.011 – marginally higher than the pooled estimation (0.010). Similarly, when estimating correct responses the coefficient more than doubles from 0.004 to 0.009. Thus, if anything, unobservable characteristics have produced a downward bias of the effect of the opponents' gender on women's performance.

Table 9 moves to scenarios three and four, focusing on *Daily Double* clues. When estimating performance, the OLS methodology again produces virtually identical coefficients. As before, incorporating fixed effects substantially increases the coefficient of

Table 8: Results from fixed effects regressions estimating whether contestant responds (columns 1 – 3) and responds correctly (columns 4 – 6). Logit regressions display marginal effects.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Answering			Answering correctly		
Estimation method:	Logit	OLS	OLS	Logit	OLS	OLS
Female	-0.005**	-0.005***		-0.002	-0.002	
	(0.002)	(0.002)		(0.002)	(0.002)	
Female × 2 male opponents	0.009***	0.010***	0.011**	0.004*	0.004*	0.009*
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.006)
Control variables[a]	yes	yes	yes	yes	yes	yes
Player fixed effects			yes			yes
# of players	8,169	8,169	8,169	8,169	8,169	8,169
% female players	39.97	39.97	39.97	39.97	39.97	39.97
N	749,433	749,433	749,433	248,052	248,052	248,052
Log lik.	-471,938			-100,660		

Notes: Standard errors clustered on the player level 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 round, a binary indicator for the *Double Jeopardy!* round, year fixed effects, and an indicator for whether the respondent has answered a clue incorrectly before.

interest (from 0.033 to 0.047), as competing against two males again improves women's performance. However, the derived coefficient is not significantly different from zero in column (3). Note that standard errors are more than doubled when fixed effects are employed, indicating a much more restrictive estimation. Thus, it is possible that the fixed effects framework does not leave enough variation to reveal potential differences along the lines of gender composition.

Table 9: Results from fixed effects regressions estimating whether contestant responds correctly to *Daily Double* clue (columns $1-3$) and the wagering amount in *Daily Double* clues (columns 4 and 5). Logit regressions display marginal effects.

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Answering correctly in DD			Wagering decision in DD	
Estimation method:	Logit	OLS	OLS	OLS	OLS
Female	-0.029*** (0.011)	-0.029*** (0.011)		-0.033*** (0.005)	
Female × 2 male opponents	0.033** (0.013)	0.033** (0.014)	0.047 (0.031)	0.037*** (0.006)	0.051*** (0.016)
Control variables[a]	yes	yes	yes	yes	yes
Player fixed effects			yes		yes
# of players	6,075	6,075	6,075	6,075	6,075
% female players	44.16	44.16	44.16	44.16	44.16
N	12,615	12,615	12,615	12,615	12,615
Log lik.	-8,064				

Notes: Standard errors clustered on the player level 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 round, a binary indicator for the *Double Jeopardy!* round, year fixed effects, and an indicator for whether the respondent has answered a clue incorrectly before.

Finally, column (4) reproduces the most complete OLS regression estimating wagering decisions. When moving to a fixed effects framework, the familiar result is recovered and women are willing to take on higher risks when competing against males. Here again, the derived coefficient gains in magnitude, this time from 0.037 to 0.051. Note that this coefficient is estimated with much more statistical precision (significant on the one percent level), even though standard errors are again elevated compared to the pooled regression (almost tripled from 0.006 to 0.016).

Overall, these results from implementing fixed effects regressions strongly confirm our baseline findings. Female *Jeopardy!* contestants perform better and take on higher risks when competing against two males, even when relying on within-player variation only. Further robustness checks are referred to the appendix Tables AII to AV.

5 Discussion

Our findings are surprising, as they stand in stark contrast to the closest predecessors. Related to risk preferences, Säve-Söderbergh and Lindquist (2016) find women to wager *less* when competing against males in the Swedish edition of *Jeopardy!*.[11] We find the opposite and it is worthwhile to compare the two setups to identify potential reasons why results are so different.

First, Säve-Söderbergh and Lindquist (2016) analyze data from Sweden, whereas our data is derived from the US.[12] Differences across countries in the effect of the opponents' gender on performance and risk-taking behavior have previously been identified. For example, Cárdenas et al. (2012) find substantial differences across Colombian and Swedish schoolgirls. Their findings indicate no gender gap in Colombia, whereas Swedish girls are, in fact, more competitive in some tasks than boys. In another study, Andersen et al. (2013) find differences between matrilineal and patriarchal societies in how females (after puberty) compete. In general, cross-country differences in the role of women have been identified across a wide range of studies, potentially owed to religion, culture, or other factors. Thus, our results may simply differ from Säve-Söderbergh and Lindquist (2016) because we analyze a US sample and they study a Swedish sample. In addition, although

[11] Other studies conform to these findings (e.g., Booth and Nolen, 2012b; Booth et al., 2014), although most of the evidence is derived for school kids or first-year college students.

[12] Säve-Söderbergh and Lindquist (2016) also discuss this explanation in their literature review.

speculative at this point, it is likely that the Swedish sample is much more homogeneous than the US sample, given the vast ethnic diversity in the US population.

Second, the US edition of *Jeopardy!*, with its long history, allows us to analyze a much larger sample than the Swedish edition. In fact, our sample is almost 15 times larger than the sample from Säve-Söderbergh and Lindquist (2016) with 8,169 versus 556 contestants. When relating the number of observations for the study of risk preferences, our sample is over 22 times larger (12,615 versus 556). In fact, Säve-Söderbergh and Lindquist (2016) are able to incorporate only 35 *Daily Double* clues where a woman competes against two men. Our analysis produces 2,527 such observations. Sample sizes from the relevant experimental studies are noticeably smaller still (see Table 1). Thus, the remarkable consistency with which we find women to compete better when paired against men may emerge in large sample sizes, as outliers may drive results in smaller samples.

Third, our rich data set allows for the inclusion of additional control variables that are absent from Säve-Söderbergh and Lindquist (2016). Most importantly, we control for the 20 most prominent question categories (e.g., STEM), but also the score of the contestant relative to their opponents, year fixed effects, and the question number in the current episode. Basic gender differences have been pointed out along several of these dimensions by other studies, such as selection into STEM categories in education, a potential learning effect in competitive environments by women and so on.

Fourth and most importantly, the large US sample allows us to estimate performance and risk attitudes in a fixed-effects framework. In other words, we are able to control for any remaining individual heterogeneity along the lines of education, any other sorts of individual *Jeopardy!*-relevant knowledge, and respective behavioral patterns. This alleviates concerns about a potential selection bias, as those women assigned to compete against two males may, a priori, simply be better at the game of *Jeopardy!*.

In summary, our analysis differs along several dimensions from the closest predecessor by Säve-Söderbergh and Lindquist (2016). The fact that results differ substantially should encourage further research into whether the opponents' gender is able to explain gender differences in performance and risk-taking.

6 Conclusions

This paper analyzes a rich database of 4,279 *Jeopardy!* episodes with 8,169 contestants to study gender differences in performance and risk preferences. Being unable to choose the gender of their opponents, this setting proves ideal to study whether the gender of females' opponents affects their behavior in a highly competitive situation with large stakes.

We study four distinct scenarios: the likelihood of responding, responding correctly, and responding correctly in particularly high-stakes situations (*Daily Double* clues); and wagering decisions in *Daily Double* situations. Along all performance dimensions, a woman fares significantly better when paired against males and the gender gap disappears. Even the substantial gender gap in risk preferences vanishes completely when a woman is competing against two males. These results are robust to the inclusion of a number of control variables and alternative estimations. In addition, these findings are confirmed when player fixed effects are included, thereby exploiting within-player variation only.

Our findings relate to several areas of policymaking. First, gender separation in education may lead to under-performance and more risk aversion among females. Second, some occupations continue to be highly gender-specific. As an example, Eckel and Füllbrunn (2015) suggest that a higher share of women in finance may alleviate risk taking. Our conclusions imply that inserting one woman into an all-male environment may not evoke risk aversion. Rather, she may adapt to males' behavior. However, integrating more women in such work environments may indeed produce more risk aversion among women. Given the setup of our study (*Jeopardy!* always features three contestants), we cannot evaluate at which exact gender ratio the differences emerge. This provides a promising area of future studies. Finally, in a more general context, our results support the idea that gender differences in performance and risk preferences depend on social environments (specifically the gender of the opposition), rather than gender-inherent particularities.

Of course, it is possible that *Jeopardy!* contestants are not necessarily representative of the entire population. Nevertheless, the strength with which our results emerge in such a large sample hints at a general importance of gender composition in explaining gender gaps in performance and risk preferences. Women may perform better and take on higher risks when competing against men.

References

Alesina, A. and Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, 53(4):898–944.

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.

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

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.

Boult, A. (2016). Single-sex schools 'put girls at huge disadvantage'. *The Telegraph, January 5, 2016.*

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

Cambridge News (2016). Girls do not suffer 'huge disadvantage' from single-sex education, says cambridge headteacher. *Cambridge News, January 6, 2016.*

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.

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

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

DoL (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.

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.

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

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

Gneezy, U. and Pietrasz, A. (2013). When half of the men are more competitive than all women. *Working Paper*.

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

Goldin, C. (2014). A grand gender convergence: Its last chapter. *The American Economic Review*, 104(4):1091–1119.

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.

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.

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

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.

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. (2015). Game, set, and match: Do women and men perform differently in competitive situations? *Journal of Economic Behavior & Organization*, 119:96–108.

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.

Lavy, V. (2013). Gender differences in market competitiveness in a real workplace: Evidence from performance-based pay tournaments among teachers. *The Economic Journal*, 123(569):540–573.

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.

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.

Lott-Lavigna, R. (2016). Society – not single-sex schooling – teaches women not to speak out. *The Guardian, January 7, 2016*.

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.

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.

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

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.

Reuben, E., Sapienza, P., and Zingales, L. (2015). Competitiveness and the gender gap among young business professionals. *IZA Discussion Paper #9446*.

Säve-Söderbergh, J. and Lindquist, G. S. (2016). Cildren 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.

Sutter, M. and Glätzle-Rützler, D. (2014). Gender differences in the willingness to compete emerge early in life and persist. *Management Science*, 61(10):2339–2354.

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

Turner, K. and de Normann, A. (2016). Dear headmaster: Single-sex or co-ed – which is better? *The Telegraph, January 7, 2016.*

Appendix

Table AI: Summary statistics of control variables by gender.

Variable	(1) Females Mean (N)	(2) Males Mean (N)	(1) = (2) p-value
Panel A: Answering (N = 749,433)			
STEM	0.056	0.056	0.416
$ value of clue	740	718	0.000***
Score	3,996	3,862	0.000***
Relative score	150	-100	0.000***
Missed clue before	0.622	0.616	0.000***
Panel B: Answering correctly (N = 248,052)			
STEM	0.057	0.056	0.557
$ value of clue	713	692	0.000***
Score	4,299	4,110	0.000***
Relative score	1,270	896	0.000***
Missed clued before	0.618	0.616	0.275
Panel C: Daily Double clues (N = 12,615)			
STEM	0.064	0.071	0.117
$ value of clue	1,045	1,013	0.000***
Score	6,121	5,770	0.000***
Relative score	3,293	2,834	0.001***
Missed clued before	0.704	0.701	0.694

Table AII: Robustness checks I. Dependent variable is whether contestant responds. Results from Logit regressions are displayed as marginal effects. Please see footnotes for description of each displayed regression.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Estimation method:	Logit	OLS	Logit	OLS	Logit	OLS	Probit
Female	-0.005**		-0.005***		-0.002		-0.005***
	(0.002)		(0.002)		(0.003)		(0.002)
Female × 2 male opponents	0.009***	0.011**	0.009***	0.011**	0.009***	0.012*	0.009***
	(0.002)	(0.005)	(0.002)	(0.005)	(0.003)	(0.006)	(0.002)
STEM	0.006**	0.007***	0.006**	0.007***	0.007*	0.009**	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
$ value of clue	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Score	0.000***	0.000***	0.000***	0.000***	0.000***	0.000*	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Question #	-0.001***	-0.001***	-0.002***	-0.001***	-0.002***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Double Jeopardy! round	-0.019***	0.005*	-0.030***	-0.002	-0.025***	0.008	-0.020***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.003)
Relative score	0.004***	-0.001***	0.004***	-0.001***	0.004***	-0.001*	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missed clue before	0.023***	-0.015***	0.021***	-0.016***	0.022***	-0.017***	0.023***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Answered clue correctly before			0.064***	0.040***			
			(0.003)	(0.003)			
Year FE & category FE (20)	yes	yes	yes	yes	yes	yes	yes
Player FE		yes		yes		yes	
N	749,433	749,433	749,433	749,433	416,625	416,625	749,433
Log lik.	-471,940	-484,298	-471,604	-484,168	-261,147	-267,883	-471,941

Notes: Standard errors clustered on the player level displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Columns (1) and (2): Displaying all control variables of column (7) in Table 4 and column (3) in Table 8. Columna (3) and (4): Including binary indicator for player having responded correctly to a clue before. Columns (5) and (6): Using only shows that have been consecutively available since January 5, 2004. Column (7): Using probit framework.

Table AIII: Robustness checks II. Dependent variable is whether contestant responds correctly. Results from Logit regressions are displayed as marginal effects. Please see footnotes for description of each displayed regression.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Estimation method:	Logit	OLS	Logit	OLS	Logit	OLS	Probit
Female	-0.002		-0.002		-0.001		-0.002
	(0.002)		(0.002)		(0.002)		(0.002)
Female × 2 male opponents	0.004*	0.009*	0.004*	0.010*	0.004	0.014**	0.004*
	(0.002)	(0.006)	(0.002)	(0.006)	(0.003)	(0.007)	(0.002)
STEM	-0.018***	-0.019***	-0.018***	-0.019***	-0.021***	-0.022***	-0.018***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.003)
$ value of clue	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Score	0.000***	0.000	0.000***	-0.000	0.000***	-0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Question #	-0.003***	-0.003***	-0.003***	-0.003***	-0.002***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Double Jeopardy round	-0.006**	-0.010***	-0.008***	-0.009**	0.011**	0.033***	-0.006**
	(0.003)	(0.004)	(0.003)	(0.004)	(0.005)	(0.006)	(0.003)
Relative score	-0.001***	-0.001***	-0.001***	-0.001***	0.000**	0.001**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missed clue before	-0.003*	0.061***	-0.004**	0.061***	-0.003	0.056***	-0.003*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Answered clue correctly before			0.014***	-0.007**			
			(0.004)	(0.003)			
Year FE & category FE (20)	yes	yes	yes	yes	yes	yes	yes
Player FE		yes		yes		yes	
N	248,052	248,052	248,052	248,052	136,620	136,620	248,052
Log lik.	-100,663	-85,512	-100,656	-85,509	-54,739	-45,894	-100,649

Notes: Standard errors clustered on the player level displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Columns (1) and (2): Displaying all control variables of column (7) in Table 4 and column (3) in Table 8. Columna (3) and (4): Including binary indicator for player having responded correctly to a clue before. Columns (5) and (6): Using only shows that have been consecutively available since January 5, 2004. Column (7): Using probit framework.

Table AIV: Robustness checks III. Dependent variable is whether contestant responds correctly to *Daily Double* clue. Results from Logit regressions are displayed as marginal effects. Please see footnotes for description of each displayed regression.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Estimation method:	Logit	OLS	Logit	OLS	Logit	OLS	Probit
Female	-0.029***		-0.028***		-0.019		-0.029***
	(0.011)		(0.011)		(0.014)		(0.011)
Female × 2 male opponents	0.033**	0.047	0.033**	0.047	0.038**	0.026	0.033**
	(0.013)	(0.031)	(0.013)	(0.031)	(0.017)	(0.040)	(0.013)
STEM	0.033*	0.034**	0.033*	0.034**	0.027	0.028	0.033*
	(0.017)	(0.017)	(0.017)	(0.017)	(0.022)	(0.021)	(0.017)
$ value of clue	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Score	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Question #	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Double Jeopardy round	-0.020	-0.020	-0.020	-0.020	-0.028	-0.031	-0.020
	(0.019)	(0.019)	(0.019)	(0.019)	(0.030)	(0.031)	(0.019)
Relative score	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Missed clue before	0.010	0.011	0.010	0.011	0.016	0.017	0.009
	(0.011)	(0.011)	(0.011)	(0.011)	(0.015)	(0.015)	(0.011)
Answered clue correctly before			-0.090	-0.060			
			(0.176)	(0.117)			
Year FE & category FE (20)	yes	yes	yes	yes	yes	yes	yes
Player FE		yes		yes		yes	
N	12,615	12,615	12,615	12,615	7,006	7,006	12,615
Log lik.	-8,064		-8,064		-4,468		-8,064

Notes: Standard errors clustered on the player level displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Columns (1) and (2): Displaying all control variables of column (7) in Table 4 and column (3) in Table 8. Columna (3) and (4): Including binary indicator for player having responded correctly to a clue before. Columns (5) and (6): Using only shows that have been consecutively available since January 5, 2004. Column (7): Using probit framework.

Table AV: Robustness checks IV. Results from OLS regressions. Dependent variable is share of maximum wager in Daily Double clue. Please see footnotes for description of each displayed regression.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.033***		-0.033***		-0.027***		-161.086***	
	(0.006)		(0.006)		(0.007)		(31.566)	
Female × 2 male opponents	0.037***	0.051***	0.037***	0.051***	0.034***	0.040*	146.705***	214.952***
	(0.007)	(0.016)	(0.007)	(0.016)	(0.009)	(0.021)	(38.404)	(82.786)
STEM	-0.004	0.010	-0.004	0.010	0.008	0.030**	26.111	66.718
	(0.008)	(0.010)	(0.008)	(0.010)	(0.011)	(0.013)	(53.709)	(67.893)
$ value of clue	-0.000*	0.000	-0.000*	0.000	-0.000	-0.000	-0.019	0.066
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.051)	(0.064)
Score	0.000	0.000	0.000	0.000	0.000**	0.000	0.232***	0.205***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.011)	(0.014)
Question #	-0.012***	-0.011***	-0.012***	-0.011***	-0.013***	-0.012***	-28.287***	-22.292***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(2.022)	(2.672)
Double Jeopardy round	-0.239***	-0.254***	-0.239***	-0.254***	-0.265***	-0.261***	-19.038	135.272*
	(0.010)	(0.012)	(0.010)	(0.013)	(0.015)	(0.020)	(51.598)	(70.222)
Relative score	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.069***	-0.072***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.006)
Missed clue before	0.031***	0.026***	0.031***	0.027***	0.035***	0.028***	34.664	-62.209
	(0.005)	(0.008)	(0.005)	(0.008)	(0.007)	(0.011)	(29.307)	(43.111)
Answered clue correctly before			-0.091	-0.072				
			(0.063)	(0.053)				
Year FE & category FE (20)	yes	yes	yes	yes	yes	yes	yes	yes
Player FE		yes		yes		yes		yes
N	12,615	12,615	12,615	12,615	7,006	7,006	12,615	12,615

Notes: Standard errors clustered on the player level displayed in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Columns (1) and (2): Displaying all control variables of column (7) in Table 4 and column (3) in Table 8. Columna (3) and (4): Including binary indicator for player having responded correctly to a clue before. Columns (5) and (6): Using only shows that have been consecutively available since January 5, 2004. Columns (7) and (8): Estimating actual wagering amount, as opposed to share of maximum.