Gender Differences in Risk Preference:

A Case Study in *Jeopardy!*

Diab Tony Eid

Presented to the Department of Economics
in partial fulfillment of the requirements
for a Bachelor of Arts degree with Honors

Harvard College
Cambridge, Massachusetts

March 19, 2020

Abstract

This paper examines high-stakes wagering on the game show *Jeopardy!* to compare risk preferences between men and women. A prior analysis on *Jeopardy!* by Andrew Metrick (1995) examines a dataset of 1,150 players to estimate a coefficient of absolute risk aversion expressed on the show with no comparison across gender. I note three shortcomings of Metrick's methodology, and I update his model to accurately compare risk preference across gender on a novel dataset consisting of 15,240 players. The first analysis of this paper finds men to wager slightly more as a fraction of their game income than women on the show, and the result is significant. However, a more refined model that takes into account utility, subjective probability, and overconfidence does not find enough evidence to suggest men and women exhibit different risk tolerances.

Acknowledgements

This thesis would not have been possible without the help of many generous friends, family members, professors, and mentors. I would first like to express my deepest gratitude to my thesis advisor, Nathaniel Hendren, and to my seminar instructor, Judd Cramer, for their patience, suggestions, and assistance over the course of several months. Thank you both for helping make this project a reality. I would also like to express thanks to all my seminar peers for their valuable comments in class, especially to Sophia Campbell who suggested looking into gender on *Jeopardy!*; to Sam Stemper and David Kane for their help with R; to Cristopher Patvakanian for being a great roommate and providing lots of moral (and Latex) support; to Yasmin Luthra for her detailed draft commments; to Emery Liu and Raymond Lin for their math help; to Jeffrey Miron for his helpful comments; to Lan Luo for his data-scraping help; and to Larry Gu for his extraordinary help with R, Latex, and the math. I would also like to thank all my teachers past and present, my parents for their love and support, and God for always being by my side.

Contents

1 Introduction

It is commonly believed that men are more risk-seeking than women (Kay 2013). Some argue this gap in risk preference explains gender disparities in occupation choice and, as a consequence, broader economic outcomes (Dohmen et al. 2011). Others believe such gaps are attributable to institutional barriers (Leythienne and Ronkowski 2018) or some other variable that differs across gender (Blau and Kahn 2000). This paper contributes to the debate by examining wagers on the game show *Jeopardy!* to compare risk profiles between men and women. I ultimately find little evidence to support the claim that risk preference differs significantly across gender.

The game show environment provides abundant data to estimate risk attitudes, but prior literature in the field focuses on calculating specific risk parameters for small samples of people with no comparison across gender. Gertner (1993), for example, examines betting on *Card Sharks* to estimate a coefficient of constant absolute risk aversion (CARA) that best explains the decision making of individuals on the show. Metrick (1995) employs a similar methodology with wagers on *Jeopardy!*. This paper builds off of Metrick's model by correcting for its major shortcomings and introducing a comparison of risk tolerance across gender.

Metrick examines situations known as "lock games." In a lock game, the leading player going into Final Jeopardy has more than double the score of the trailer. A lock game leader has a sure victory so long as he or she does not bet more than an amount that would put him or her at risk of being caught by the trailer. With victory assured, the leader is only worried about how much money he or she will take home. Assuming individuals follow exponential utility, wagers in lock games will express some level of constant absolute risk aversion. Metrick assesses wagers over a subset of games to calculate a single CARA coefficient without comparison across gender. His CARA coefficient estimate is 6.6×10^{-5} with a standard error of 5.6×10^{-5}.

The Metrick model serves as the foundation of this paper, but it has three main flaws. First, the nature of Final Jeopardy prevents players from making negative wagers, which introduces bias into the risk parameter estimation. Second, the model relies on a logistic regression that forces the intercept term to be zero, which the data might violate. Third, the model does not disentangle risk-seeking behavior from overconfidence on the show.[1]

While the first flaw prevents meaningful conclusions regarding the level of risk tolerance a specific group of people might display on the show, it does not prevent relative comparisons of risk tolerance across gender. The second and third flaw can also be mitigated with an updated model and greater

[1] This last shortcoming is mentioned by Metrick on page 245 of his paper.

access to data. This paper adjusts the model, applies it to a dataset about 14 times as large, and introduces cross-gender risk comparison.

The rest of the paper is organized as follows. Section 2 presents a review of the relevant literature and an explanation of the rules of *Jeopardy!*. Section 3 presents data and summary statistics for all games in *Jeopardy!* history. Section 4 presents the analyses and Section 5 is the conclusion.

2 Background

2.1 Literature Review

Prior research mostly supports the idea that women are less risk-seeking than men, although the question is by no means settled. Watson and McNaughton (2007) find women to choose more conservative retirement investment strategies than men. Investing decisions more generally also exhibit this gap (See Sunden and Surrette 1998; Olsen and Cox 1998; and Hinz, McCarthy, and Turner 1997 for examples). Moreover, Zinkhan and Karande (1991) find women less likely to take business risks than men.

Lab studies do not paint a clear picture on whether men are more risk-seeking than women. Schubert et al. (1999) find that, of five treatments, women are more risk-averse than men in two, more risk-seeking in one, and not significantly different from men in the other two. In another study by Holt and Laury (2002), women are found to be more risk-averse than men when the stakes are low (a few dollars), but when the stakes are high (greater than ten dollars), there is not enough evidence to conclude women are more or less risk-seeking than men. Finally, Eckel and Grossman (2002) find women to be more risk-averse than men when faced with a series of gamble choices.

The main limitation of lab studies is the magnitude of the stakes: risking tens of dollars is incommensurate with the high-stakes decisions individuals must face in their day-to-day lives. Game shows, however, typically involve large gambles. For example, the average amount a player can wager in Final Jeopardy is about $9,346, and the mean wager is $4,949.[2] These gambles provide researchers opportunities to calculate specific levels of risk tolerance. In the literature, researchers apply some form of a utility function to these situations and fit the data to their specification.

A common choice for game show modeling is exponential utility, which contains a single risk parameter: the coefficient of constant absolute risk aversion (CARA). Gertner (1993) estimates a CARA coefficient lower bound for bets on *Card Sharks* of 3.1×10^{-4} with a standard error of

[2]These figures come directly from the dataset described in Section 3 and do not include non-regular season play.

1.7×10^{-5}. Metrick's (1995) analysis of wagers on *Jeopardy!* found a CARA coefficient of 6.6×10^{-5} with a standard error of 5.6×10^{-5}, which was not significantly different from risk neutrality. A paper on *Lingo* by Beetsma and Schotman (2001), on the other hand, estimates a strong risk-averse CARA coefficient of 0.12 with a standard error of 0.01. That paper also found constant relative risk aversion (CRRA) to explain gambles roughly as well as the CARA specification.

Some papers in the game show environment advocate utilizing a more flexible utility function, such as the expo-power utility function introduced by Saha (1993).[3] The main issue with a more flexible utility specification is the danger of over-fitting. Fitting multiple parameters to a series of observed gambles to minimize a loss function will lose some interpretative value as one adds more parameters to the model. Moreover, Hartog J. and Jonker (2000) find that, for financial wealth up to $100,000, CARA is a reasonably good description of risk attitudes. In light of this, exponential utility is the preferred specification of this paper.

2.2 The Rules of *Jeopardy!*

This section provides a brief explanation of the rules of *Jeopardy!* as of March 12th, 2020. *Jeopardy!* is a trivia game with three players. The goal of the game is to finish with the most money. A player collects money by answering questions correctly.[4] Questions are initially hidden on a game board behind certain dollar values. The larger the dollar value, the tougher the question. At the top of each column on the game board is the category for all the questions in that column. The category is visible to all players.

The game is divided into three rounds. In the first round, the "Jeopardy Round," the returning champion picks one of the squares on the board, e.g. "I'll take Famous Economists for $400." Then, the question is shown and read by the host to all three players. When the question is done being read, a green light appears around the game board, indicating to all three players that they are allowed to buzz in. If a player buzzes in before the green light turns on, that person is prevented from buzzing for half a second. The first person to buzz in is the person who gets the first chance to answer the question. That person has five seconds to do this. If he or she answers incorrectly or runs out of time, the dollar value of the question is deducted from his or her current score. If he or she answers correctly, he or she earns the dollar value of that clue. In the case of an incorrect answer or time running out, the green light turns back on and the other players have the chance to

[3] See Post et al. (2008) for one application of this model on the game show *Deal or No Deal*.

[4] Technically, the show presents players with "answers" and they have to give the "questions." For example, the answer "This Economist won the Nobel Prize in 1976 for his work on consumption analysis, monetary history, and the complexity of stabilization policy" would merit the question "Who is Milton Friedman?" I ignore this technicality throughout the paper.

buzz in. The question is live until someone gets it right, all three get it wrong, or no one buzzes in. This last situation is called a "Triple Stumper." Players begin the game with zero dollars and scores can become negative.

The person that ultimately answers correctly has "control of the board," meaning he or she gets to pick the next question. The process repeats until all squares have been picked or until time for the Jeopardy Round runs out. Hiding under one of the squares on the board in the Jeopardy Round is the "Daily Double." If a player hits a Daily Double, he or she has the opportunity to wager some amount of his or her money on the question. If he or she answers correctly, he or she wins their wager. If he or she answers incorrectly, that amount is deducted from his or her score. The category is known to the player before wagering, but the specific question is not. The second round, the "Double Jeopardy Round," is the exact same as the first round except all dollar values are doubled and two Daily Doubles are hidden on the board.

The third round is "Final Jeopardy." In this round, all three players are shown one category for one final question. Each must make a wager up to the amount he or she has. No player can see the wager of any other player. Then, the question is revealed, and each player has 30 seconds to write down his or her answer. A correct answer gives the reward of his or her wager. An incorrect answer deducts the player's wager from his or her score.

The player with the most money after Final Jeopardy wins both the game and the amount of money he or she earned. He or she also gets to return the next day to play again as the returning champion. The other two players win a pre-determined amount of money: $2,000 and $1,000 for second and third place, respectively.

Lock games, which are the focus of the estimation methods in this paper, occur when the leading player going into Final Jeopardy has more than double the score of the trailer. In these games, the leader has a sure victory so long as he or she does not bet more than an amount that would put him or her at risk of being caught by the trailer.

2.3 Rule Changes

2.3.1 Lock Games

The frequent *Jeopardy!* viewer might notice that the definition of a lock game has evolved over time due to rule changes on the show. Before November 2014, if a game ended in a draw, the players in the draw would all return to play the next day. After November 2014, if a game ends in a draw, Alex Trebek, the host, asks a tiebreaker question and only one person emerges victoriously. As a consequence, a game in which the leader has exactly double the score of the trailer going into Final Jeopardy is considered a lock before the rule change, but not after. The frequency of such events is small: of all 5,174 games observed, there are only 37 in which the leader has exactly double the score of the trailer going into Final Jeopardy. The analyses of this paper will therefore define a lock game to be a situation in which the leader has strictly more than double the score of the trailer going into Final Jeopardy.

2.3.2 Returning Champions

In today's version of the show, a player can win and come back the next day for an unlimited amount of time. Prior to September 8, 2003, however, winners were not allowed to return after five wins. Brad Rutter, considered one of the best players of all time, had his time on *Jeopardy!* before the cap was removed. This rule change has no effect on most analyses presented in this paper, except for the "Streak" variable in general summary statistics. This variable is capped at a value of five for games before September 8, 2003, potentially skewing the average.

3 Data

3.1 All *Jeopardy!* Games

Metrick examines data from 393 *Jeopardy!* games between October 1989 and January 1992. This translates to 1,150 Final Jeopardy decisions made by around 1,000 individuals. The discrepancy between the number of decisions and the number of individuals is due to the fact that champions return to play another game. He does not have any demographic information, but he notes that the average *Jeopardy!* player is likely wealthier and better educated than the average person (an assumption that I see no reason to doubt, given the game's difficulty).

The first task of this paper is to build a dataset that consists of every game in *Jeopardy!* history by scraping the J!Archive. The archive has information about many different elements of a specific game: who got which questions right, which questions went unanswered, how much players wagered for Final Jeopardy, and more.[5] All data analysis is done in R (R Core Team 2013). Tables are created using the stargazer package (Hlavac 2018) and figures are made with ggplot (Wickham 2016).

The scraped dataset consists of 5,080 regular season *Jeopardy!* games and 15,240 individual

[5]The J!Archive is an online repository available at www.j-archive.com. The scraped dataset is available upon request. See Appendix for more information.

Final Jeopardy decisions.[6] Summary statistics are provided in Table 1. Note that a single player is represented multiple times when assessing Final Jeopardy decisions, because winners return to play the next day and must make another wager for Final Jeopardy. For example, Ken Jennings made 74 *Jeopardy!* appearances, all of which are included in the summary statistics of Table 1.

Table 1: Summary Statistics for all *Jeopardy!* Players

Statistic	Mean	St. Dev.	Min	Median	Max
Score Going into Final Jeopardy ($)	9,345.67	6,529.98	−6,800	8,200	72,600
Wager ($)	4,948.53	4,246.71	0	4,000	60,013
Correct	0.46	0.50	0	0	1
Final Score ($)	9,520.32	9,731.90	−6,800	7,106	131,127
Final Coryat ($)	9,325.95	5,698.57	−6,800	8,400	39,200
Male	0.59	0.49	0.00	1.00	1.00
Streak	2.47	5.34	1	1	74

Notes: This table displays summary statistics for 15,240 players on the game show *Jeopardy!*. "Score Going into Final Jeopardy" represents a player's score immediately before the Final Jeopardy round, "Wager" represents a player's Final Jeopardy wager, "Correct" is a dummy variable indicating whether or not a player correctly answered the Final Jeopardy question, "Final Coryat" represents a player's score if all wagering in the game is disregarded, "Male" is a dummy variable indicating if a player is male or not, and "Streak" indicates how many games the returning champion has won. These data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012.

We can see that stakes on the show are high. The average amount a player can wager on Final Jeopardy is about $9,346, and the mean wager is $4,949. Final Jeopardy also proves to be challenging for the average contestant: players answer correctly only 46% of the time. This follows the *Jeopardy!* community's intuition that Final Jeopardy tends to be as difficult as the fourth row of clues on the board (the fifth row is considered the hardest).

We also see some interesting results on the extreme ends of each variable. James Holzhauer's record for most money won before Final Jeopardy appears in the dataset ($72,600), as well as Ken Jennings' longest win streak (74 games). A careful reader might worry about the negative value appearing in the dataset as the minimum for "Final Score." In regular season play, it is indeed possible to finish the game with a negative amount of money, but that player is not allowed to participate in Final Jeopardy.

With respect to gender, the J!Archive does not directly provide information on player gender identification. The archive does, however, give each player's name. Utilizing the gender package in R, I examine the frequency of a player's name in Social Security Administration data between 1930 and 2012 (Mullen 2019). In my analysis, I use the "ssa" method from the package, which draws

[6] These numbers exclude tournament and other non-regular *Jeopardy!* play.

from Social Security Administration Data in the United States between 1930 and 2012. When given a name, the package outputs the frequency that name appeared as male in the Social Security data between 1930 and 2012. I then take this frequency and classify it as the probability that the given name is male. To be categorized as male in the analyses of this paper, one's probability of being male must exceed 95%. That same threshold is also set for women. As a result, about 1,741 observations are dropped from gender analyses in this paper, and I am able to assign a gender to 13,499 players. Of the 15,240 total players in my dataset, about 52% are men, 36% are women, and 11% are unable to be categorized. The male skew in candidate selection could be the result of a number of factors, including, but not limited to, the age of the show, gender bias in candidate selection, and preference differences.

Tables 2 and 3 present summary statistics for men and women, respectively. Men tend to have about $1,900 more than women going into Final Jeopardy. Men also tend to wager more than women, though the gap is smaller at about $500. Men answer Final Jeopardy correctly 48% of the time, while that number for women is 42%. On average, men have higher Final Scores and Coryat Scores than women, with gaps of about $2,500 and $1,500 respectively. These data suggest a gender gap in ability on the show. Assessing the causes or ramifications of these statistics is beyond the scope of this paper.

Table 2: Summary Statistics for Male *Jeopardy!* Players

Statistic	Mean	St. Dev.	Min	Median	Max
Score Going into Final Jeopardy ($)	10,120.21	7,150.25	−5,000	8,800	72,600
Wager ($)	5,130.50	4,495.64	0	4,100	60,013
Correct	0.48	0.50	0	0	1
Final Score ($)	10,518.34	10,614.97	−5,000	8,194.5	131,127
Final Coryat($)	9,970.89	6,033.02	−5,000	9,000	39,200
Streak	2.66	6.04	1	1	74

Notes: This table displays summary statistics for 7,944 male players on the game show *Jeopardy!*. "Score Going into Final Jeopardy" represents a player's score immediately before the Final Jeopardy round, "Wager" represents a player's Final Jeopardy wager, "Correct" is a dummy variable indicating whether or not a player correctly answered the Final Jeopardy question, "Final Coryat" represents a player's score if all wagering in the game is disregarded, and "Streak" indicates how many games the returning champion has won. These data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012.

Table 3: Summary Statistics for Female *Jeopardy!* Players

Statistic	Mean	St. Dev.	Min	Median	Max
Score Going into Final Jeopardy ($)	8,245.18	5,516.81	−6,800	7,400	42,200
Wager ($)	4,696.18	3,899.74	0	4,000	21,600
Correct	0.42	0.49	0	0	1
Final Score ($)	8,066.64	8,326.68	−6,800	5,800	46,801
Final Coryat ($)	8,404.62	5,120.21	−6,800	7,600	30,400
Streak	2.24	4.46	1	1	74

Notes: This table displays summary statistics for 5,555 female players on the game show *Jeopardy!*. "Score Going into Final Jeopardy" represents a player's score immediately before the Final Jeopardy round, "Wager" represents a player's Final Jeopardy wager, "Correct" is a dummy variable indicating whether or not a player correctly answered the Final Jeopardy question, "Final Coryat" represents a player's score if all wagering in the game is disregarded, and "Streak" indicates how many games the returning champion has won. These data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012.

3.2 Lock Games

In a lock game, the leading player going into Final Jeopardy has more than double the score of the trailer. A lock game leader has a sure victory so long as he or she does not bet more than an amount that would put him or her at risk of being caught by the trailer. Therefore, up to the maximum bet, the leader is wagering some amount of money on his or her ability to answer the question correctly. If we assign some subjective probability to the player answering correctly, his or her wager would then reveal some level of risk tolerance. This estimation is the focus of Section 4 of this paper.

Table 4 presents summary statistics for lock games. Notice that, in the average lock game, the leader has up to about $4,500 to wager. This is the maximum value the leader can afford to lose in case the trailer doubles his or her money. Also, lock game leaders have a five percentage point higher success rate when answering Final Jeopardy than the average player. Lock game leaders also have higher Coryat Scores, Double Jeopardy Scores, and Final Scores when compared to the average player. This is to be expected, as anyone who achieves a lock game on *Jeopardy!* is typically highly skilled at the game.

Tables 5 and 6 present summary statistics for male and female lock game leaders.[7] The gender gap observed for general *Jeopardy!* games persists in lock games as well. Of the 7,944 male players observed, 912 (11%) are lock game leaders. Of the 5,555 female players observed, 275 (5%) are lock game leaders. Male leaders tend to have $3,250 more than female leaders going into Final Jeopardy. Male leaders also tend to wager more, with a gap of about $1,700. Lastly, male leaders have a

[7]These leaders are not necessarily unique individuals; a returning champion is included as a new player for each game he or she plays.

Table 4: Summary Statistics for Lock Game Situations on *Jeopardy!*

Statistic	Mean	St. Dev.	Min	Median	Max
Leader's Score ($)	18,528.10	9,437.72	2,000	17,200	72,600
Trailer's Score ($)	6,050.66	3,121.78	−1,400	5,600	18,800
Leader's Wager ($)	3,515.57	5,572.65	0	2,000	60,013
Leader Got It?	0.51	0.50	0	1	1
Trailer Got It?	0.41	0.49	0	0	1
Streak	4.25	9.35	1	2	74

Notes: This table displays summary statistics for 1,334 lock games on the game show *Jeopardy!*. Lock games are situations in which the leader going into Final Jeopardy has more than double the score of the trailer. "Leader's Score" is the leader's score going into Final Jeopardy, "Trailer's Score" is the trailer's score going into Final Jeopardy, "Leader's Wager" is the leader's wager, "Got It?" is a dummy variable indicating if the respective player actually got the question correct, and "Streak" indicates how many games the returning champion has won. These data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012.

success rate of 54%, while that number for female leaders is 46%.

The number of lock games is substantially lower than the number of *Jeopardy!* games: about 8.8% of all games are lock games, which corresponds to 1,334 observations. After removing lock games where the leader's name provides no clear indication of gender, we are left with 1,187 lock games (7.8% of all games). Over time, more lock games are bound to occur, and the analyses of this paper can be repeated on a much larger dataset.

Table 5: Summary Statistics for Male Lock Game Leaders on *Jeopardy!*

Statistic	Mean	St. Dev.	Min	Median	Max
Leader's Score ($)	19,530.00	10,246.40	3,300	18,200	72,600
Trailer's Score ($)	6,180.47	3,181.06	−1,400	5,650	18,000
Leader's Wager ($)	4,061.30	6,468.10	0	2,000	60,013
Leader Got It?	0.54	0.50	0	1	1
Trailer Got It?	0.41	0.49	0	0	1
Streak	5.31	11.17	1	2	74

Notes: This table displays summary statistics for 912 male lock game leaders on the game show *Jeopardy!*. Lock games are situations in which the leader going into Final Jeopardy has more than double the score of the trailer. "Leader's Score" is the leader's score going into Final Jeopardy, "Trailer's Score" is the trailer's score going into Final Jeopardy, "Leader's Wager" is the leader's wager, "Got It?" is a dummy variable indicating if the respective player actually got the question correct, and "Streak" indicates how many games the returning champion has won. These data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned by utilizing Social Security Administration data between 1930 and 2012.

Table 6: Summary Statistics for Female Lock Game Leaders on *Jeopardy!*

Statistic	Mean	St. Dev.	Min	Median	Max
Leader's Score ($)	16,280.18	7,086.67	2,000	15,600	42,200
Trailer's Score ($)	5,754.91	2,992.40	−200	5,500	18,800
Leader's Wager ($)	2,325.25	2,423.85	0	1,600	15,000
Leader Got It?	0.46	0.50	0	0	1
Trailer Got It?	0.41	0.49	0	0	1
Streak	1.99	1.53	1	1	10

Notes: This table displays summary statistics for 275 female lock game leaders on the game show *Jeopardy!*. Lock games are situations in which the leader going into Final Jeopardy has more than double the score of the trailer. "Leader's Score" is the leader's score going into Final Jeopardy, "Trailer's Score" is the trailer's score going into Final Jeopardy, "Leader's Wager" is the leader's wager, "Got It?" is a dummy variable indicating if the respective player actually got the question correct, and "Streak" indicates how many games the returning champion has won. These data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012.

4 Analyses

4.1 Model Definitions

I begin by defining variables that will be referenced across all models presented in this paper. Let the amount of money a player has going into Final Jeopardy (FJ) be denoted x, a player's wager in FJ be denoted y, a player's subjective probability of getting FJ correct be denoted p^s, a player's CARA coefficient be denoted α, and the present value of a player's lifetime wealth be denoted W.

A player is indexed according to his or her rank after Double Jeopardy, denoted i, and the specific *Jeopardy!* game he or she is in, denoted j. For example, a player's wager on show number 234 who is in first place after Double Jeopardy is represented as $y_{1,234}$, and the second place player's wager after Double Jeopardy on that same episode is denoted $y_{2,234}$. For simplicity, the index j will only be used for the regressions of this paper.

Since players must wager no more than what they earned before FJ and no less than zero, we have the constraint $0 \leq y_i \leq x_i$. Also, we can write a lock game like so: $x_1 > 2x_2$. Under this model's assumptions, a lock game leader who wagers an amount that no longer ensures victory is considered irrational. So, for a rational lock game leader, we have the additional constraint $y_1 < x_1 - 2x_2 - 1$. In our data, of the 1,334 lock games, 133 have irrational lock game leaders. These individuals are filtered out from all lock game analyses.

4.2 Linear Model

Before estimating player CARA coefficients under an exponential utility specification, I run a linear regression to get a general sense of risk preference expressed on the show. The dependent variable of this model is a lock game leader's wager as a fraction of the maximum possible bet. As described in the prior section, the rational leader's maximum possible bet is $x_1 - 2x_2 - 1$. I regress this fraction on a gender dummy variable and introduce controls for ability. The regression equation is given by Equation (1). The ability vector, $\vec{A}$, contains five ability controls: the number of questions a player answered correctly over the game, the number of questions a player answered incorrectly over the game, a player's Coryat Score (a score that disregards all wagering during the game), a player's score after Double Jeopardy, and a player's Final Score. A full description of these variables can be found in the Appendix.

$$\frac{y_{1,j}}{x_{1,j} - 2x_{2,j} - 1} = \beta_0 + \beta_1 \times \text{Male}_j + \vec{\beta_j} \cdot \vec{A_j} + \epsilon_j \tag{1}$$

Since some names do not provide a clear enough indication of gender, they are dropped from the regression. Of 1,334 lock games, 139 are dropped due to this ambiguity and 133 are dropped due to irrationality ($y_1 > x_1 - 2x_2 - 1$). We are left with 1,062 observations.

The results are displayed in Table 7. Under Regressions (1), being male is associated with wagering slightly more as a fraction of one's allowed wager—around five percentage points more—and the effect is significant at the five percent level. After introducing ability controls for number of questions answered correctly and incorrectly, that number increases to 7.5 percentage points and the effect is significant at the ten percent level. After introducing all five controls for ability, the difference settles to six percentage points with significance at the five percent level. The R^2 values for all five regressions are very small, indicating that gender accounts for only a small fraction of the variation observed in wagers.

These results suggest men are willing to risk more of their money than women, even after controlling for ability. However, this analysis by no means paints a full picture of risk tolerance differences between both groups. These regressions do not fully account for the diminishing marginal utility of money. They also control for ability rather than the subjective probability of answering Final Jeopardy correctly. The analysis can be further complicated by introducing an exponential utility model and estimating a coefficient of absolute risk aversion that best fits the model to the data. This is the task of the next section.

Table 7: Risk Estimation from Final Jeopardy Wagers

	Dependent variable:					
	Wager as a Fraction of Maximum Possible Bet					
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.049**	0.069***	0.075***	0.066***	0.059**	0.057**
	(0.023)	(0.024)	(0.024)	(0.023)	(0.023)	(0.023)
Correct		−0.005***	−0.008***	−0.002	−0.003	−0.004*
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Incorrect			−0.009***	−0.010***	−0.008***	−0.008***
			(0.002)	(0.002)	(0.002)	(0.002)
Coryat				−0.00001***	−0.00002***	−0.00002***
				(0.00000)	(0.00000)	(0.00000)
Score After DJ					0.00001***	−0.00000
					(0.00000)	(0.00000)
Final Score						0.00001***
						(0.00000)
Constant	0.579***	0.700***	0.879***	0.870***	0.890***	0.917***
	(0.020)	(0.044)	(0.065)	(0.064)	(0.063)	(0.063)
Observations	1,062	1,062	1,062	1,062	1,062	1,062
R^2	0.004	0.013	0.026	0.049	0.073	0.095
Adjusted R^2	0.003	0.011	0.023	0.045	0.069	0.090

Notes: This table presents linear regressions that estimate the risk profile of male and female lock game leaders on the game show *Jeopardy!*. Lock games are situations in which the leader going into Final Jeopardy has more than double the score of the trailer. The "Maximum Possible Bet" for a lock game leader is the one that ensures victory even if the trailer doubles his or her money. "Male" is a dummy variable indicating if a player is male. "Correct" and "Incorrect" represent how many questions the player answered correctly and incorrectly, respectively. "Coryat" refers to a player's Coryat Score, which is a score that disregards all wagering during the game. "Score After DJ" is a player's score after the Double Jeopardy Round and "Final Score" is a player's score at the end of the game. The regressions suggest that, after controlling for ability, being male is associated with wagering slightly more (about six percentage points) as a fraction of one's maximum possible bet. Game data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012. Standard errors are in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

4.3 Metrick Model

I now replicate Metrick's (1995) model of lock game situations in *Jeopardy!* and note three of its shortcomings. Assume players follow exponential utility with constant absolute risk aversion (CARA) as given by Equation (2). Under this function, the sign of α, the coefficient of absolute risk aversion, determines if an individual is risk-seeking ($\alpha < 0$), risk-neutral ($\alpha = 0$), or risk-averse ($\alpha > 0$). Note that this equation omits the divisor common to exponential utility. Since players ultimately solve the maximization problem given by Equation (3), this assumption has no impact on the model's results.

$$U(x) = 1 - e^{-\alpha x} \tag{2}$$

The leader in a lock game situation solves:

$$\max_{y_1} [p_1^s U(x_1 + y_1 + W_1) + (1 - p_1^s) U(x_1 - y_1 + W_1)] \quad \text{s.t.} \quad 0 \leq y_1 \leq x_1 - 2x_2 - 1 \tag{3}$$

The optimal wager is therefore given by:

$$y_1^* = \min[\frac{\ln\left(\frac{p_1^s}{1-p_1^s}\right)}{2\alpha_1}, x_1 - 2x_2 - 1] \tag{4}$$

The intuition of the optimal wager is straightforward: the leader follows exponential utility and makes a wager based on his or her subjective probability of getting FJ correct. However, due to the nature of *Jeopardy!*, if this optimal value exceeds the limit imposed by the possibility of second place doubling his or her money, the leader will simply bet $x_1 - 2x_2 - 1$ so as to ensure victory while maximizing utility. This type of wager occurs exactly 129 times in lock game history.

If the optimal value does not exceed the limit imposed by the trailer, then the optimal wager is given by Equation (5). In these situations, observation of y_1^* and p_1^s on the show allows for an estimation α_1, the CARA coefficient.

$$y_1^* = \frac{\ln\left(\frac{p_1^s}{1-p_1^s}\right)}{2\alpha_1} \tag{5}$$

I now explain the first flaw of this model. Metrick notes that, with uncensored knowledge of every player's p_1^s and y_1^*, he can read out α_1 for every individual through Equation (5). Let us assume one has access to that information. If the optimal wager (y_1^*) exceeds the limit imposed by the constraint in Equation (2), then one cannot observe a risk parameter for this individual. Such a wager occurs 129 times out of 1,347 lock games, which accounts for 0.096% of all lock games.

Similarly, *Jeopardy!* does not allow negative wagers, so risk-averse individuals with $\alpha_1 > 0;\ p_1^s < 0.5$ and risk-seeking individuals with $\alpha_1 < 0;\ p_1^s > 0.5$ will want to wager some negative amount but will be forced by the nature of the game to wager zero. So, for players who wager zero, we cannot calculate a precise CARA coefficient.[8] The zero wager in a lock game situation occurs 115 times, which accounts for 0.085% of all lock games.

The infrequency of the two aforementioned problematic wagers suggests that omitting them from analyses will have little effect on overall risk estimation. However, given the general assumption of risk aversion in economics, omitting zero wagers from the analyses could skew the CARA coefficient mean to be more risk-seeking than it actually is. Given this possibility, using the model to make conclusions about risk tolerance for a certain population, as Metrick does, is a bit of a leap. While CARA coefficient estimation under Metrick's model may be flawed, relative comparisons of CARA coefficients across groups is still justified, assuming both groups are affected similarly by the bias. Thus, this paper omits problematic wagers from all analyses and refrains from making conclusions about the level of risk tolerance for any one particular group. Instead, I draw relative risk comparisons across gender.

Now comes the issue of calculating a subjective probability p_1^s for every lock game leader. Metrick does not do this. Instead, he rearranges Equation (5) and notices that it follows the form of a logistic regression.

$$\ln\left(\frac{p_1^s}{1-p_1^s}\right) = 2y_1^*\alpha_1 \tag{6}$$

Equation (6) is the form of a logistic regression with an intercept term that is forced to be zero. "Answering Final Jeopardy Correctly" is the binary dependent variable, and two times a player's wager is the independent variable. The coefficient of interest is α_1. In Metrick's original analysis, he found the intercept term of a free logistic regression to not be significantly different from 0. This, he argued, allowed him to keep the model's assumption of a zero intercept. This assumption is data-dependent. Under a different dataset, the assumption of a zero intercept might be violated. If one allows equation (6) to have a free intercept, then the optimal wager y_1^* is given by:

[8] Of course, the idea of allowing negative wagers on *Jeopardy!* is silly, but in the context of this model, "Answering Final Jeopardy Correctly" is treated as a random event, and individuals desiring negative wagers for random events over which they have no control is reasonable.

$$\ln\left(\frac{p_1^s}{1-p_1^s}\right) = 2\alpha_1 y_1^* + \beta_1 \tag{7}$$

Here we find the second flaw in Metrick's methodology. It can be mathematically shown that, under classical assumptions, no closed form utility function exists such that Equation (7) is the solution to the initial maximization problem laid out in Equation (3) when $\beta_1 \neq 0$. The proof is as follows.

Proof. We want to show that there does not exist a continuous function $U(x;\alpha,\beta)$ with the properties $U'(x) > 0; x, \alpha, \beta \in \mathbb{R}$ such that, for all $x_1 > 0, W_1 > 0$, and $p_1^s \in [0,1]$, the solution to

$$\max_{y_1}[p_1^s U(x_1 + y_1 + W_1) + (1-p_1^s)U(x_1 - y_1 + W_1)] \tag{a}$$

yields

$$\ln\left(\frac{p_1^s}{1-p_1^s}\right) = 2\alpha_1 y_1^* + \beta_1 \tag{b}$$

where β_1 is nonzero. Assume there does exist some utility function that satisfies this. Let $A = x_1 + W_1$. The first order condition brought about by the maximization problem is:

$$\frac{p_1^s}{1-p_1^s} = \frac{U'(A-y_1^*)}{U'(A+y_1^*)} \tag{c}$$

Let $L = A - y_1^*$ and $c = 2y_1^*$. Then we can rewrite the first order condition like so:

$$\frac{p_1^s}{1-p_1^s} = \frac{U'(L)}{U'(L+c)} \tag{d}$$

Now, let $f(x) = U'(x)$. Considering this must be true for all values of $A > 0$, we rewrite again:

$$\frac{p_1^s}{1-p_1^s} = \frac{f(x)}{f(x+c)} \tag{e}$$

$$\implies f(x+c) = \left(\frac{1-p_1^s}{p_1^s}\right) f(x)$$

By substituting in Equation (b):

$$\frac{f(x)}{f(x+c)} = e^{\alpha_1 c + \beta} \tag{f}$$

$$\implies \frac{f(x)}{f(x+2c)} = e^{2\alpha_1 c + \beta}$$

Also, note that, by Equation (e), we have:

$$f(x+2c) = \left(\frac{1-p_1^s}{p_1^s}\right)^2 f(x) \tag{g}$$

$$\implies \frac{f(x)}{f(x+2c)} = e^{2\alpha_1 c + 2\beta}$$

$$\implies e^{2\alpha_1 c + \beta} = e^{2\alpha_1 c + 2\beta}$$

$$\implies \beta = 2\beta$$

$$\implies \beta = 0$$

Therefore, Metrick's model can only be interpreted if the intercept term of a free logistic regression is not significantly different from zero. QED.

I now run Metrick's logistic regression to see if the data violates the zero intercept assumption. Table 8 shows the results of this regression broken up by gender. The first three columns follow the logistic regression given by Equation (7)—they regress "Answering Final Jeopardy Correctly" on two times a lock game leader's wager. Regression (1) includes rational lock game leaders and omits the problematic wagers mentioned earlier. Regressions (2) and (3) are limited to identifiable men and women, respectively. Regression (4) is a difference-in-differences analysis that compares the male and female coefficients from Regressions (2) and (3). Notice that the observation count for Regression (4) is smaller than that of Regression (1). This is due to the fact that Regression (4) compares players that could be categorized as men and women: ambiguously named individuals are not included. In Regression (1), however, ambiguously named players are included.

None of the coefficients are significantly different from zero, meaning the Metrick model retains some interpretative value here. Regression (4) indicates that the CARA coefficient mean is not significantly different between men and women. While the intercept terms are not significantly different from zero, interpreting Metrick's model would be impossible had these terms been nonzero, a possibility that is more likely given a larger dataset.

The third problem with Metrick's model has to do with overconfidence. Let us assume that players on the show exhibit overconfidence, i.e., their subjective probability of answering FJ correctly is greater than their true probability. Then, by Equation (6), CARA coefficients will appear more risk-seeking than they actually are. Under-confidence, on the other hand, would make CARA coefficients appear more risk-averse than they actually are. Metrick mentions overconfidence as a possible explanation behind the relatively low CARA coefficient he estimated for his sample.

The three shortcomings mentioned thus far can be alleviated with some updates to the model. First, given the bias caused by omitting players with zero wagers and maximum wagers, we can focus solely on inter-group comparison of risk tolerance, assuming the bias equally affects both groups. Second, we can move away from a reliance on single-parameter estimation via logistic regression by allowing the probability of getting FJ correct (p_1^s) to vary by individual. This allows us to calculate α_1 for every lock game leader in the show's history. Third, we can adjust the model to include an overconfidence parameter and see how its inclusion affects our results. These alterations are discussed in the following section.

Table 8: CARA Coefficient Estimates via Logistic Regression for Lock Game Leaders on *Jeopardy!*

	Dependent variable:			
	Answering Final Jeopardy Correctly			
	All	Men	Women	All
	(1)	(2)	(3)	(4)
Wager Doubled	0.00003***	0.00004***	−0.00001	−0.00001
	(0.00001)	(0.00001)	(0.00003)	(0.00003)
Male				0.146
				(0.208)
Wager Doubled × Male				0.00004
				(0.00003)
Constant	0.014	0.113	−0.033	−0.033
	(0.076)	(0.093)	(0.186)	(0.186)
Observations	1,201	819	243	1,062
Log Likelihood	−808.796	−539.429	−168.317	−707.746
Akaike Inf. Crit.	1,621.591	1,082.858	340.634	1,423.493

Notes: This table displays estimates for a coefficient of constant absolute risk aversion (CARA) that best explains the decision making of lock game leaders on *Jeopardy!*. Lock games are situations in which the leading player going into Final Jeopardy has more than double the score of the trailer. In such games, the leader's optimal wager under an exponential utility model follows the form of a logistic regression where "Answering Final Jeopardy Correctly" is the binary dependent variable and "Wager Doubled" is the independent variable. The coefficient on "Wager Doubled" is the CARA coefficient that best explains the wagers of the respective population. A lower CARA coefficient corresponds to greater risk-seeking behavior. The last column presents a difference-in-differences analysis to compare the coefficients expressed by men and women in Regressions (2) and (3). The difference between men and women is not statistically significant. Game data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012. Standard errors are in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

4.4 Updated Metrick Model

Metrick did not have access to probability information for each game. Hence, he was required to utilize logistic regression across all games to estimate one risk parameter for the behavior of the entire population. Thanks to the abundance of information on the J!Archive, I can adjust Metrick's model and calculate player specific subjective probabilities (p_1^s) and CARA coefficients (α_1). To impute a p_1^s to each player, the adjusted model regresses the binary variable "Answering Final Jeopardy Correctly" on five in-game variables across all *Jeopardy!* games to predict a player's true probability (p_i) of answering FJ correctly.[9] The main assumption behind such a regression is that to assess an individual's true probability of getting FJ correct, one should see how that person and his or her opponents fared in the game and then examine FJ success rates for similar types of games over the history of the show. This becomes the true probability p_i that a player has of getting FJ correctly. Let us return to the optimal wager equation under the Metrick Model:

$$y_1^* = \frac{\ln\left(\frac{p_1^s}{1-p_1^s}\right)}{2\alpha_1} \tag{5}$$

Now, assuming (i) a player's subjective probability is equal to the probability predicted by the logistic regression ($p_i = p_i^s$) and (ii) players make an optimal wager following this model, the leader's unadjusted CARA coefficient (α_1) is given by rearranging Equation (5):

$$\alpha_1 = \frac{\ln\left(\frac{p_1}{1-p_1}\right)}{2y_1^*}$$

The CARA coefficient is unadjusted because it does not take into account overconfidence: it assumes that $p_i = p_i^s$. Perhaps players have some level of overconfidence or under-confidence ($p_i \neq p_i^s$). In the model, we can define confidence bias (denoted c) as the difference between a player's subjective probability (p^s) and true probability (p). Thus, we can write: $p_1^s = p_1 + c_1$. Notice that when $c_1 < 0$, we have under-confidence, and when $c_1 > 0$, we have overconfidence. Substituting this into Equation (5) yields:

$$y_1^* = \frac{\ln\left(\frac{p_1+c_1}{1-p_1-c_1}\right)}{2\alpha_1} \tag{8}$$

Since Equation (8) has two unknowns, α_1 and c_1 for each player, we cannot make a calculation of α_1 for each individual. Instead, we can estimate two values $\hat{\alpha_1}$ and $\hat{c_1}$ that best explain the wagers

[9]See Appendix for an in-depth explanation of this regression.

made by the population. The particular method of fit I use is non-linear least squares estimation. If we index every lock game leader by j and look at n lock games, we can represent the least-squares problem by Equation (9).

$$\min_{\hat{\alpha_1},\hat{c_1}} \left[\sum_{j=1}^{n} \left(\frac{\ln\left(\frac{p_j+\hat{c_1}}{1-p_j-\hat{c_1}}\right)}{2\hat{\alpha_1}} - y_j^* \right)^2 \right] \tag{9}$$

The non-linear fit analysis chooses $\hat{c_1}$ and $\hat{\alpha_1}$ that minimizes the sum of square errors in Equation (9) for men, and then also does this optimization separately for women. Results are displayed in Table 9. Women tend to overestimate their probability of getting FJ correct by 7.42 percentage points, while that number for men is 3.03. Each error is significant at the one percent level.

Table 9: Least Squares Estimation of CARA Coefficients and Overconfidence by Gender

Parameter	Estimate	St. Error	Statistic	p-value
Male CARA Coefficient	0.000170	1.37e-05	12.42	1.37e-32
Male Overconfidence	0.030253	5.86e-03	5.16	3.03e-07
Female CARA Coefficient	0.000291	4.56e-05	6.39	8.54e-10
Female Overconfidence	0.074214	1.77e-02	4.20	3.80e-05

Notes: This table displays the results of two non-linear least squares estimations that fit an overconfidence measure and a coefficient of constant absolute risk aversion (CARA) to the Final Jeopardy (FJ) wagers of lock game leaders on the game show *Jeopardy!*. Lock games are situations in which the leader going into FJ has more than double the score of the trailer. Assuming lock game leaders optimize over an exponential utility function, a CARA coefficient and overconfidence level can be fit to the wagering decisions observed on the show. According to the table, men tend to overestimate their probability of getting FJ correct by 3.03 percentage points, while women tend to overestimate their probability by 7.42 percentage points. Men also exhibit a more risk-seeking CARA coefficient, but if one incorporates overconfidence into a model that allows CARA coefficients to vary by individual, the difference in risk tolerance becomes insignificant. See Table 12 for the results of that specification. Game data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012.

A direct comparison of $\hat{\alpha_1}$ across gender is possible at this point. Indeed, the male CARA coefficient is lower than the female CARA coefficient and the result is statistically significant. However, this eliminates the model's allowing of α_1 to vary by individual. So, I return to Equation (8) and add $\hat{c_1}$ to each player's predicted p_1. Now, each player has a subjective probability (p_1^s) and we can read out an adjusted CARA coefficient for each player. Recall that the unadjusted CARA coefficient is the CARA coefficient for each player assuming $p_1 = p_1^s$ (no overconfidence).

Summary statistics for unadjusted and adjusted CARA coefficients are displayed Tables 10 and 11, respectively. Notice that in Table 10, the mean unadjusted CARA coefficient for men and women is negative, which indicates risk-seeking behavior. However, in Table 11, we see that, after adjusting for overconfidence, the mean becomes positive (risk aversion). Recall that CARA coefficient estimation in this context has bias due to the fact that players cannot wager a negative amount or more than the maximum amount. Assuming the bias is equal across men and women, however, we can make relative comparisons across gender.

Table 10: Unadjusted CARA Coefficient Summary Statistics by Gender

Group	Mean	St. Dev.	Min	Median	Max	N
Men	-1.08e-04	0.0147	-0.1773	6.64e-05	0.29318	819
Women	-7.21e-05	0.0034	-0.0494	3.26e-05	0.00822	243

Notes: This table shows summary statistics for uadjusted coefficients of absolute risk aversion expressed by lock game leaders on the game show *Jeopardy!*. Lock games are situations in which the leader going into Final Jeopardy (FJ) has more than double the score of the trailer. Under an exponential utility specification, a lock game leader's coefficient of constant absolute risk aversion (CARA) is a function of his or her FJ wager and his or her subjective probability of answering FJ correctly. Subjective probabilities are imputed to each player via a logistic regression that regresses "Answering FJ Correctly" on five in-game variables. Unadjusted CARA coefficients do not take into account overconfidence: they assume subjective probabilities are equal to the probabilities generated by the logistic model. Game data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned using Social Security Administration data between 1930 and 2012.

Table 11: Adjusted CARA Coefficient Summary Statistics by Gender

Group	Mean	St. Dev.	Min	Median	Max	N
Men	0.000239	0.01554	-0.1518	9.03e-05	0.3604	819
Women	0.000334	0.00189	-0.0188	1.19e-04	0.0117	243

Notes: This table shows summary statistics for adjusted coefficients of absolute risk aversion expressed by lock game leaders on the game show *Jeopardy!*. Lock games are situations in which the leader going into Final Jeopardy (FJ) has more than double the score of the trailer. Under an exponential utility specification, a lock game leader's coefficient of constant absolute risk aversion (CARA) is a function of his or her FJ wager and his or her subjective probability of answering FJ correctly. Subjective probabilities are imputed to each player via a logistic regression that regresses "Answering FJ Correctly" on five in-game variables. Adjusted CARA coefficients account for overconfidence by running non-linear least squares estimation of confidence bias across gender and adding those estimates to the probabilities generated by the logistic model. Game data are scraped from the J!Archive, accessible at www.j-archive.com. Gender is assigned using Social Security Administration data between 1930 and 2012.

To assess if the differences observed in Tables 10 and 11 are significant across gender, I run two linear regressions. The results are shown in in Table 12. In Regression (1), being male is associated with a slightly lower unadjusted CARA coefficient, indicating more risk-seeking behavior. When taking into account overconfidence, the gap between the average male and female CARA coefficient grows. However, for both unadjusted and adjusted analyses, the difference between men and women is not significant. Moreover, extremely low R^2 and Adjusted R^2 values suggest gender does not do much to explain the variation in CARA coefficients across individuals.

This methodology is not perfect. For one, overconfidence and under-confidence is not allowed to vary by individual. The only confidence variation permitted by the model is variation across gender lines. There are simply not enough within-individual observations to fit some level of overconfidence to every lock game leader. Still, the separation of confidence bias and true probability is a meaningful first step toward a better understanding of the relationship between wagers and risk preference on *Jeopardy!*.

Table 12: Linear Regression of CARA Coefficients on Gender in *Jeopardy!*

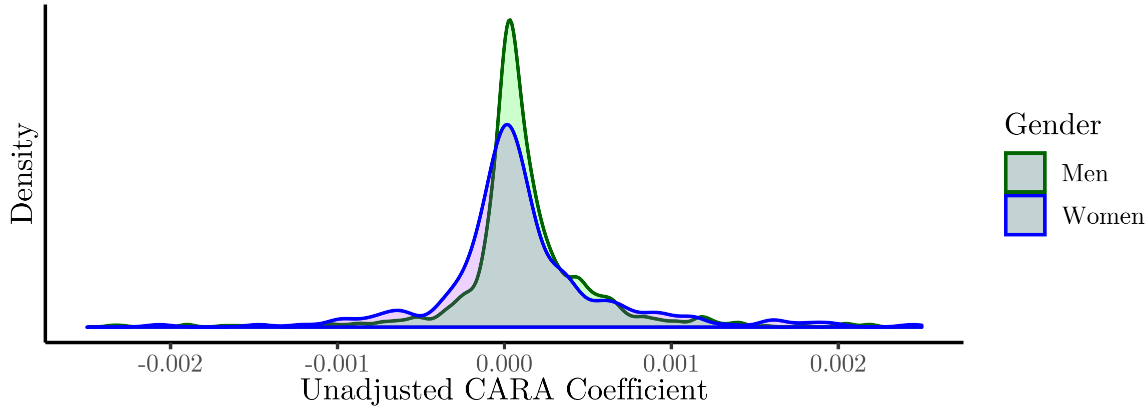
	Dependent variable:	
	Unadjusted CARA Coefficient	Adjusted CARA Coefficient
	(1)	(2)
Male	−0.00004	−0.0001
	(0.001)	(0.001)
Constant	−0.0001	0.0003
	(0.001)	(0.001)
Observations	1,062	1,062
R^2	0.00000	0.00001
Adjusted R^2	−0.001	−0.001

Notes: This table shows the relationship between a lock game leader's gender and level of risk tolerance expressed on the game show *Jeopardy!*. Lock games are situations in which the leader going into Final Jeopardy (FJ) has more than double the score of the trailer. Under an exponential utility specification, a lock game leader's coefficient of constant absolute risk aversion (CARA) is a function of their wager and their subjective probability of answering FJ correctly. Subjective probabilities are imputed to each player via a logistic regression that regresses "Answering FJ correctly" on five in-game variables. "Adjusted CARA Coefficient" accounts for overconfidence by running non-linear least squares estimation of confidence bias across gender and adding those estimates to the probabilities generated by the logistic model. "Unadjusted CARA Coefficient" does not account for overconfidence. "Male" is a dummy variable indicating if a player is male or not. There is not enough evidence to suggest that the average male coefficient of absolute risk aversion is significantly different from the average female coefficient. Game data are scraped from the J!Archive, accessible at www.j-archive.com. Gender information is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012. Standard errors are in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Figure 1 presents a visual comparison of unadjusted CARA coefficient densities by gender. Since these coefficients are unadjusted, they assume a player's subjective probability is equal to the probability generated by the logistic model. For both men and women, the distribution has a normal shape with a slight left skew. Negative outliers push the average unadjusted CARA coefficient below zero for both groups, as indicated by Table 10. There is also slightly more bunching around the mean for men than for women. Overall, however, both genders exhibit similarly shaped unadjusted CARA coefficient distributions when not accounting for overconfidence.

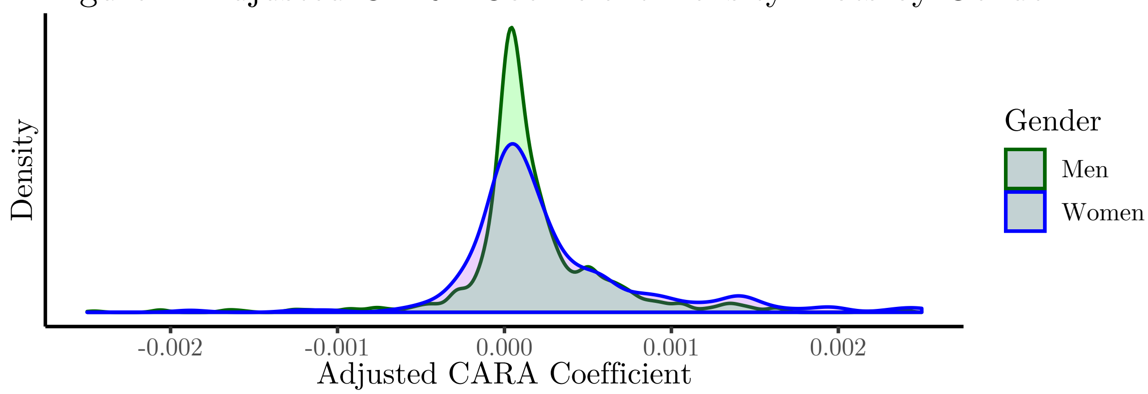
Figure 2 plots density against adjusted CARA coefficients that account for overconfidence. As evidenced by the figures in Table 11, the means for both men and women are pushed above zero. The density plots for men and women have slightly more of a left skew than before, and both maximum and minimum values have increased, indicating a rightward shift of the distribution. This makes sense, given that overconfidence leads to higher subjective probabilities and, as a result, higher CARA coefficients. The female distribution remains slightly wider than the male distribution.

Figure 1: Unadjusted CARA Coefficient Density Plots by Gender



Notes: This graph plots the densities of unadjusted CARA coefficients revealed through Final Jeopardy (FJ) wagers by lock game leaders on Jeopardy. Lock games are situations in which the leader going into Final Jeopardy has more than double the score of the trailer. Under an exponential utility specification, a lock game leader's coefficient of constant absolute risk aversion (CARA) is a function of his or her FJ wager and his or her subjective probability of answering FJ correctly. Subjective probabilities are imputed to each player via a logistic regression that regresses "Answering FJ Correctly" on five in-game variables. Unadjusted CARA coefficients do not take into account overconfidence: they assume subjective probabilities are equal to the probabilities generated by the logistic model. Game data are scraped from the J!Archive, accessible at www.j-archive.com. Gender information is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012.

Figure 2: Adjusted CARA Coefficient Density Plots by Gender



Notes: This graph plots the densities of adjusted CARA coefficients revealed through Final Jeopardy (FJ) wagers by lock game leaders on Jeopardy. Lock games are situations in which the leader going into Final Jeopardy has more than double the score of the trailer. Under an exponential utility specification, a lock game leader's coefficient of constant absolute risk aversion (CARA) is a function of his or her FJ wager and his or her subjective probability of answering FJ correctly. Subjective probabilities are imputed to each player via a logistic regression that regresses "Answering FJ Correctly" on five in-game variables. Adjusted CARA coefficients account for overconfidence by running non-linear least squares estimation of confidence bias across gender and adding those estimates to the probabilities generated by the logistic model. Game data are scraped from the J!Archive, accessible at www.j-archive.com. Gender information is assigned by pulling name frequencies from Social Security Administration data between 1930 and 2012.

4.5 Robustness Check

Perhaps there is some error in each $p_{1,j}$ predicted by the logistic regression under the updated model. Under a different specification, one might see a difference in risk preferences emerge. Recall that the model regresses the dummy variable "Answering Final Jeopardy Correctly" on five in-game variables across all games in the show's history to predict a $p_{1,j}$ for each player. This section checks for the robustness of this method by seeing how the results change if the logistic regression included only four or fewer in-game variables.

In Table 12, two linear regressions allowed us to compare the difference in means of adjusted and unadjusted CARA coefficients across gender. I now repeat the analysis entirely but under five different probability model specifications that impute a p_1 for each player. The first model regresses "Answering Final Jeopardy Correctly" on "Number of Questions Answered Correctly During the Game" and uses the results to predict a probability of answering FJ correctly for each player. The second adds "Number of Questions Answered Incorrectly" to the regression; the third adds "Coryat Score"; the fourth adds "Score After Double Jeopardy Round"; and the fifth adds "Place After Double Jeopardy Round." The fifth specification is the one originally utilized in Section 4.4 of this paper.

Table 13: Probability Model Robustness Check

Model	Diff. in Means (Men-Women)	St. Error	t-score	p-value
(1)	3.77e-04	0.000320	1.1781	0.239
(2)	3.75e-05	0.000466	0.0804	0.936
(3)	-1.28e-04	0.000560	-0.2283	0.819
(4)	-1.78e-04	0.000553	-0.3226	0.747
(5)	-9.48e-05	0.000556	-0.1705	0.865

Notes: This table shows how the difference in risk tolerance estimates between male and female lock game leaders on *Jeopardy!* changes under different probability models. Lock games are situations in which the leader going into Final Jeopardy (FJ) has more than double the score of the trailer. Under an exponential utility specification, a lock game leader's coefficient of constant absolute risk aversion (CARA) is a function of his or her wager and his or her subjective probability of answering FJ correctly. Subjective probabilities are imputed to each player via five different logistic regressions in this robustness check. The "Diff. in Means" is the difference between the average adjusted CARA coefficient for men and women. Adjusted CARA coefficients account for overconfidence by running non-linear least squares estimation of confidence bias across gender and adding those estimates to the probabilities generated by the logistic models. Game data are scraped from the J!Archive, accessible at www.j-archive.com. Gender information is assigned using Social Security Administration data between 1930 and 2012.

Table 13 presents the results of the robustness check. The mean CARA coefficient differences for this table are all adjusted for overconfidence. It is worth noting that each "Diff. in Means" value is equivalent to the coefficient on a male dummy variable in Regression (2) of Table 12. For example, the coefficient on "Male" in Regression (2) of Table 12 is equivalent to the difference presented in Model (5) of Table 13. Under all five probability model specifications in Table 13, not a single difference in means t-test is statistically significant; all p-values are greater than ten percent. Therefore, there is not enough evidence to conclude that men are more or less risk-seeking than women.

Still, there are many ways one could predict the true probability of answering FJ correctly. One could go deeper into the details of the game. For example, the number of difficult questions a player answered correctly could be weighted more than the number of easy questions a player answered correctly. It is therefore likely that none of these models captures the true probability of answering FJ correctly, but the persistence of no difference in risk tolerance across all models does suggest that men and women exhibit similar risk tendencies.

5 Conclusion

This paper builds off a model introduced by Andrew Metrick (1995) to compare risk preferences between men and women on the high-stakes game show *Jeopardy!*. In an exact replication of Metrick's analysis on a much larger dataset, no significant difference is found between CARA coefficient estimates for men and women. This result holds even after adjusting for three limitations of Metrick's methodology.

The initial analysis of this paper regresses Final Jeopardy wagers as a fraction of maximum possible bet on gender. After controlling for ability, I find men to be slightly more risk-seeking under this specification, and the result is statistically significant. Then, I complicate the model by introducing Metrick's exponential utility logistic regression to estimate a CARA coefficient for men and women. The difference between these two estimates is non-significant. However, Metrick's model has three main flaws: (i) it does not allow CARA coefficient estimation for players who wager zero or the maximum possible bet, (ii) it forces the intercept term of a logistic regression to be zero, and (iii) it does not account for overconfidence that might be exhibited by the players.

To address these concerns, I build a logistic probability model to impute a probability of answering Final Jeopardy correctly for every player on the show. This allows for a direct CARA coefficient calculation for every lock game leader. Then, I use non-linear least squares estimation to fit an overconfidence error for both genders and calculate adjusted CARA coefficients. No significant gender difference is found, and that finding is robust across five different probability model specifications.

The updated methodology is not perfect. It fails to allow confidence errors to vary by individual, leaving room for further research into the relationship between risk preference and overconfidence on the show.

There are multiple avenues for further research on this topic. One might doubt that the exponential utility model presented is an accurate reflection of human decision making, especially because the optimal wager generated by it is not dependent on lifetime wealth. A more flexible utility model could be applied to *Jeopardy!*, but one would have to either estimate lifetime wealth or fit some value to the data. A paper by Post et al. (2008) does this kind of analysis on the game show *Deal or No Deal*.

Another topic for further research is exploration into the ability gap observed on the show. Why is it that men tend to perform better than women? The dataset used for this paper could be useful to researchers exploring that topic and other cross-gender comparisons. New episodes of *Jeopardy!* air every single day, and as the J!Archive grows larger, there is more potential for large scale investigation into numerous economic questions.

Appendix

Description of Scraped Dataset

I scrape data for this paper off of the J!Archive, an online repository of every game in the show's history. The J!Archive is accessible at www.j-archive.com.

I first directly scrape nine variables for every Jeopardy game in the show's history: 6,329 games as of August 10, 2019. 85 observations are dropped from every analysis in this paper for having NA values in every game variable, and another 1,164 observations are dropped for being from tournament play or some other type of non-regular season game. This leaves us with 5080 games to analyze. The nine scraped variables for each game are defined below. The dataset is available upon request.

1. **Show Number**
   - Every Jeopardy game on the J!Archive has a unique show number.
2. **Date**
   - The date the game was aired.
3. **Game Information**
   - This variable has basic information on the episode, such as if it was a tournament game or if it had a famous player like Ken Jennings (who won 74 regular season games in a row).
4. **Player Description**
   - This variable has each player's full name and a brief description that mentions occupation and area of residency. For returning champions, this variable contains that player's streak and how much money they have collected thus far.
5. **Coryat Score**
   - A player's Coryat Score is a player's score if all wagering is disregarded. In the Coryat Score, there is no penalty for forced incorrect responses on Daily Doubles, but correct responses on Daily Doubles earn only the natural values of the clues, and any gain or loss from the Final Jeopardy Round is ignored.
6. **Score After Double Jeopardy Round**
   - This is a player's score after the Double Jeopardy Round but before the Final Jeopardy Round.

7. **Final Score**
    - This is a player's score after the game has concluded.
8. **Number of Questions Answered Correctly**
    - This is the number of questions each player answered correctly during the game.
9. **Number of Questions Answered Incorrectly**
    - This is the number of questions each player answered incorrectly during the game. This includes triple stumpers, which occur when a question goes unanswered.

The dataset is manipulated using R to create new variables and run the analyses of this paper (R Core Team 2013). Graphs and tables of this paper are provided by the ggplot and stargazer packages (Wickham 2016; Hlavac 2018). All coding work is available upon request.

Predicting Probabilities in the Updated Metrick Model

To impute a probability $p_{1,j}$ for every lock game leader in the show's history, I run a logistic regression with "Answering Final Jeopardy Correctly" as the dependent variable and 15 game statistics as the independent variables. Then, I use the results of that regression to directly predict a player's chance of answering Final Jeopardy correctly.

The 15 independent variables can be divided into five groups of variables for all three players in the game. A list of the variables and their descriptions can be found below.

1. **Number of Questions Answered Correctly**
   - This is the number of questions each player got right during the game.
2. **Number of Questions Answered Incorrectly**
   - This is the sum of the number of questions each player answered incorrectly during the game and the number of triple stumpers in the game.
3. **Coryat Score**
   - According to the J!Archive, a player's Coryat Score is a player's score if all wagering is disregarded. In the Coryat score, there is no penalty for forced incorrect responses on Daily Doubles, but correct responses on Daily Doubles earn only the natural values of the clues, and any gain or loss from the Final Jeopardy Round is ignored.
4. **Score After Double Jeopardy Round**
   - This is a player's score after the Double Jeopardy Round (or before the Final Jeopardy Round).
5. **Place After Double Jeopardy Round**
   - This is a player's place after the Double Jeopardy Round (or before the Final Jeopardy Round).

Model (1) in the robustness check corresponds to a logistic model with only the first variable in the above list for each player. Model (2) adds the second variable, and so on. Model (5) includes all five variables; I use this model for the main analysis in Section 4.4 of this paper. The results of Model (5) are displayed in Table 14.

Sources

Beetsma, Roel M. W.J., and Peter C. Schotman. 2001. "Measuring Risk Attitudes in a Natural Experiment: Data from the Television Game Show Lingo." *The Economic Journal* 111: 821–48.

Blau, Francine D., and Lawrence M. Kahn. 2000. "Gender Differences in Pay." *Journal of Economic Perspectives* 14 (4): 75–99.

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

Eckel, Catherine C., and Philip J. Grossman. 2002. "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk." *Evolution and Human Behavior* 23 (4): 281–95.

Gertner, Robert. 1993. "Game Shows and Economic Behavior: Risk-Taking on "Card Sharks"." *The Quarterly Journal of Economics* 108 (2): 507–21.

Hartog J., Ferrer-i-Carbonell A., and N. Jonker. 2000. "On a Simple Measure of Risk Aversion." University of Amsterdam.

Hinz, Richard P., David D. McCarthy, and John A. Turner. 1997. "Are Women Conservative Investors? Gender Differences in Participant Directed Pension Investments." Pension Research Council Working Papers. Wharton School Pension Research Council, University of Pennsylvania.

Hlavac, Marek. 2018. *Stargazer: Well-Formatted Regression and Summary Statistics Tables.* Bratislava, Slovakia: Central European Labour Studies Institute (CELSI).

Holt, Charles A., and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92 (5): 1644–55.

Kay, John. 2013. "Is It Better to Play It Safe or to Place Bets That Risk Bankruptcy?" *The Financial Times*, December.

Leythienne, Denis, and Piotr Ronkowski. 2018. "A Decomposition of the Unadjusted Gender Pay Gap Using Structure of Earnings Survey Data." *Eurostat Statistical Working Papers.*

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

Mullen, Lincoln. 2019. *Gender: Predict Gender from Names Using Historical Data.*

Olsen, Robert A., and Constance M. Cox. 1998. "Gender Differences in the Allocation of Assets in Retirement Savings Plans." *American Economic Review* 88 (2): 207–11.

Post, Thierry, Martin J. van den Assem, Guido Baltussen, and Richard H. Thaler. 2008. "Deal or No Deal? Decision Making Under Risk in a Large-Payoff Game Show." *The American Economic*

Review 98 (1): 38–71.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Saha, Atanu. 1993. "Expo-Power Utility: A 'Flexible' Form for Absolute and Relative Risk Aversion." *American Journal of Agricultural Economics* 75 (4): 905–13.

Schubert, Renate, Martin Brown, Matthias Gysler, and Hans Wolfgang Brachinger. 1999. "Financial Decision-Making: Are Women Really More Risk-Averse?" *American Economic Review* 89 (2): 381–85.

Sunden, Annika E., and Brian J. Surrette. 1998. "Gender Differences in the Allocation of Assets in Retirement Savings Plans." *American Economic Review* 88 (2): 207–11.

Watson, John, and Mark McNaughton. 2007. "Gender Differences in Risk Aversion and Expected Retirement Benefits." *Financial Analysts Journal* 63 (4): 52–62.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York.

Zinkhan, George M., and Kiran W. Karande. 1991. "Cultural and Gender Differences in Risk-Taking Behavior Among American and Spanish Decision Makers." *The Journal of Social Psychology* 131 (5): 741–42.