

IZA DP No. 9612

**Women Do Not Play Their Aces:  
The Consequences of Shying Away**

Jörg Claussen  
Eszter Czipor  
Mirjam van Praag

December 2015

# Women Do Not Play Their Aces: The Consequences of Shying Away

**Jörg Claussen**

*Copenhagen Business School*

**Eszter Czibor**

*University of Amsterdam*

**Mirjam van Praag**

*Copenhagen Business School,  
University of Amsterdam and IZA*

Discussion Paper No. 9612  
December 2015

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **Women Do Not Play Their Aces: The Consequences of Shying Away\***

The underrepresentation of women at the top of hierarchies is often explained by gender differences in preferences. We find support for this claim by analyzing a large dataset from an online card game community, a stylized yet natural setting characterized by self-selection into an uncertain, competitive and male-dominated environment. We observe gender differences in playing behavior consistent with women being more averse towards risk and competition. Moreover, we demonstrate how “shying away” makes female players less successful: despite no gender gap in playing skills, women accumulate lower scores than men due to their relative avoidance of risky and competitive situations.

JEL Classification: D03, J24, M52

Keywords: gender, preferences, risk, competition, performance

Corresponding author:

Mirjam van Praag  
Copenhagen Business School  
Kilevej 14a  
2000, Frederiksberg  
Denmark  
E-mail: [mvp.ino@cbs.dk](mailto:mvp.ino@cbs.dk)

---

\* We are grateful for *sauspiel.de* for making the dataset analyzed in this study available for the purpose of our research. We thank the participants at the Incentives and Behavior Change workshop, the MBEES-2015 conference and at seminars at Sussex University and Carlos III Madrid University for their insightful comments. We are particularly grateful for the helpful suggestions of Thomas Buser, Tommaso Ciarli, Catherine Eckel, Uri Gneezy and Randolph Sloof.

Women are missing from the top of organizational hierarchies around the world. As of January 2015, a mere 4.6% of CEO positions at S&P 500 companies are held by women (Catalyst Org., 2015) and 21.9% of national parliamentarians are female (UN Women, 2015). Interestingly, even women who self-select into prestigious and competitive education follow different career trajectories than men: observing high-potential graduates from elite MBA programs, Carter and Silva (2010) find that women lag behind men in advancement and compensation starting from their first jobs. According to the European Commission (2012), the pattern is no different in academia: while there was a complete gender balance at the PhD level, only one fifth of full professors were women in 2010, and the progress towards gender equality has been slow.

A popular explanation for the absence of women in leadership positions concerns gender differences in attitudes towards risk and competition. Survey evidence from companies suggests that female managers' careers are stalled because they enter promotion contests less often than men do (Barsh and Yee, 2012; Institute for Leadership and Management, 2011; Sandberg, 2013). Lloyds TSB Bank found, for instance, that although female employees are more likely than men to meet or exceed performance expectations, they tend not to apply for promotion (Desvaux et al., 2008). Such anecdotal evidence is often quoted to support the claim that women's less positive attitudes towards risk and competition not only keep them from entering uncertain and competitive environments but continue to hinder their chances of advancement even after they have sorted into such settings. Observational data from (competitive) companies rarely allow assessing the causal impact of "shying away" on the career of women who have self-selected into such firms.<sup>1</sup> This is partly due to a lack of objective performance measures but also due to confounding factors such as potential discrimination against women (Altonji and Blank, 1999) or gender differences in actual experience and labor force interruptions (Kim and Polachek, 1994).

We therefore use a very large set of naturally occurring data from a simple and abstract setting. We analyze the behavior of registered users in an online card game community. The game has some appealing features resembling the environment of interest in various aspects: there is real, strategic and repeated interaction (with feedback) between male and female players who can encounter each other as opponents or as teammates over several rounds. Uniquely, the data allow us to make a distinction between what we call 'selection' and 'playing' behavior: we observe players' choices regarding the level of risk and competition they prefer to bear in each round, and also their subsequent playing performance (i.e. the tricks they win and the points they collect) in the resulting tournaments. Players' performance is measured exactly and depends on skill, strategic behavior and luck. Since there are no real monetary incentives involved, we conjecture that members of the card game community are highly intrinsically motivated. While we acknowledge possible concerns related to the external validity of the setting, our study allows an internally valid assessment of the impact of gender differences in risky and competitive choices on achievement.

---

<sup>1</sup>A notable exception is the study of Card et al. (2015) who decompose the impact of firm-specific pay differentials on the gender wage gap into sorting and bargaining effects. Using data from Portugal, they show that women only receive about 90% of the firm-specific premiums men collect.

We demonstrate that even women who self-select into a competitive environment subsequently refrain from initiating contests they could win. Women do not only initiate fewer contests than men, they are also less likely to play “Solo” (instead of in a team) or increase the stakes of the game. Multiple rounds of repetition and feedback about past performance do not change this behavior. As a consequence, female players accumulate lower scores and appear less competent than their male competitors, despite no actual skill difference in the playing stage (that we also demonstrate). We find that the performance of female players is at least as high as males’ once their different ‘selection’ choices have been taken into account. Finally, we argue that women’s different choices in the selection stage likely result from gender differences in preferences.

The topic of gender differences in preferences and their potential impact on labor market outcomes has received considerable attention among economic scholars over the past decade. Several studies have found that women indeed differ from men along multiple dimensions, such as attitudes towards risk and competition (see for overviews Croson and Gneezy (2009), Niederle and Vesterlund (2011), Bertrand (2011) and Azmat and Petrongolo (2014)). In particular, females are shown to be more risk averse (Charness and Gneezy, 2012; Eckel and Grossman, 2008), less likely to enter a tournament (Niederle and Vesterlund, 2007) and less responsive to competitive incentives (Gneezy et al., 2003) than males. While the existence of a gender gap in preferences has been robustly established, its measured magnitude and importance is context-sensitive. For instance, the gender gap in competitiveness is much larger when the tournament is between individuals rather than between teams (Dargnies, 2012; Healy and Pate, 2011).<sup>2</sup>

Evidence for a gender gap in risk attitudes and competitiveness comes predominantly from laboratory experiments, and the actual relevance of this gap for field outcomes has been scantily researched (Bertrand, 2011). A few recent studies have used laboratory measures of behavioral preferences to explain real life choices. Competitiveness is indeed shown to be a strong predictor for outcomes such as secondary school students’ choices of academic tracks (Buser et al., 2014) or the decision to take a selective high school entrance exam (Zhang, 2013). Flory et al. (2014) demonstrate in a field experiment that the share of female job applicants decreases with the competitiveness of the offered compensation scheme. These studies directly link gender differences in competitive preferences to occupational segregation in real life.

This study complements the above-listed papers by analyzing in a natural environment the consequences of less risky/competitive choices for the outcomes of females, given their participation in a competitive game. Our results support previous findings that females tolerate less competition and risk than men. Our main contribution is that we show using naturally occurring data that even skilled women who self-select into a merit-based, discrimination-free but competitive environment will end up lagging behind men because they are less inclined to take risk and to initiate competition. While prior results have often been obtained in the lab or in

---

<sup>2</sup>Further relevant factors affecting the observed gender gap in competitiveness include the gender composition of the group of opponents (Gneezy et al., 2003; Gneezy and Rustichini, 2004; Ivanova-Stenzel and Kübler, 2011) or of the competing teams (Delfgaauw et al., 2013) and the nature of the task at hand (Grosse et al., 2014; Wieland and Sarin, 2012). Considering risk aversion, Filippin and Crosetto (2014) show that the gender gap is sensitive to the method of elicitation. Adams and Funk (2012) find in a large, unincentivized survey that female directors are actually more risk loving than their male counterparts.

smaller sized field experiments, ours has been obtained in more natural circumstances using a large and diverse sample of people (from various age groups and backgrounds). The realistic (longer term) game environment allowed us, for instance, to reveal the substantial consequences over time of seemingly innocent and small gender differences in behavior. This would not have been possible in a lab experiment. Our study highlights the importance of research on gender differences in preferences by presenting evidence that females’ differential selection choices that impede their success are to a large extent driven by their attitudes towards competition and risk.

We argue that our results carry relevant implications for organizations that strive for a gender balance in their population of employees, including the top of the hierarchy. For instance, firms often use systems that rely, more or less explicitly, on employees entering contests, such as self-nomination for promotions or employee-initiated negotiation for pay raises. These firms might have to mentor female employees to take a more pro-active or competitive approach. Alternatively, they might design mechanisms where self-selection and self-nomination are less important in the promotion process.

The rest of the study is structured as follows. In Section 2, we provide an overview of the context and our dataset. Section 3 presents and discusses our results. Section 4 concludes.

## 1 Context and data

### 1.1 The game

We use data from an online community for playing *Schafkopf*: a popular, traditional Bavarian card game. The game is known in other regions and countries as well (though with minor variations), e.g. *Doppelkopf*, *Skat* or *Sheepshead*. *Schafkopf* is a zero-sum game, played by four participants at a (virtual) table, using the unique German/Bavarian deck of cards.<sup>3</sup> Each game consists of a selection stage in which players announce their willingness to initiate a game and have the option to raise stakes, and the actual playing stage during which all distributed cards are played out trick by trick. Points collected in the playing stage determine the winner(s) of the game. The way points map into earnings (‘cents’) depends on the chosen stakes. The aim of the game is to gain as many cents as possible since the sum of the cents collected constitutes a player’s score.

The game begins with the dealer (rotating one position clockwise each round) distributing four cards to each player. The players then evaluate the strength of their first four cards and decide whether to double the stakes of the game by knocking on the table. Afterwards, the remaining four cards are distributed to each player. Players can then take the offensive role by actively initiating a game, or the passive role by playing as a partner/opponent in a game initiated by someone else at the table. Players initiating a game have a choice between three game types: the standard two-against-two-players *Sauspiel* game and the more risky and competitive one-

---

<sup>3</sup>The Bavarian deck has eight different values (in increasing rank: 7, 8, 9, 10, Jack (Unter), Queen (Ober), King (König) and Ace) in four different suits (in decreasing rank: acorn, grass, heart, bell) each, with the Queens, Jacks and heart cards being trumps.

against-three *Wenz* and *Solo* game types.<sup>4</sup> Figure A1 in the Appendix shows a screenshot from the selection stage in the online game: each player is asked if they want to initiate a *Sauspiel*, a *Wenz* or a *Solo* game, or if they prefer not to initiate a game at all (i.e. to “pass”). Passing does not result in dropping out of the round: as long as at least one person initiates a game, all four players at the table will join as partners or opponents. The highest announced game type will be played (*Solo* being the highest game type, followed by *Wenz* and *Sauspiel*) and if multiple players want to initiate the same game type, the player closest to the dealer will be given priority. If nobody initiates a game, the cards are reshuffled and a new round begins.

The game is played in eight tricks and every player has to contribute one card to each trick. Before playing the first trick, players in the opponent role are allowed to give a “Contra”, which doubles the stakes in the game. The player sitting to the left of the dealer starts the first trick and subsequent tricks are started by the winner of the last trick. The other players then all have to lay down a card in clockwise order.<sup>5</sup> The player who contributes the card with the highest rank wins the trick.<sup>6</sup>

To determine the winner of the game, each card is allocated a point value, with a total of 120 points for the whole card deck.<sup>7</sup> To win the game, the initiator (together with his/her partner in case of a *Sauspiel* game) needs to win tricks worth at least 61 points in total. Therefore it is possible to win five out of eight tricks but still lose the game if these tricks do not contain enough points. Each player of the losing party has to pay 10 cents to the winner(s) when losing a *Sauspiel* game and 50 cents when losing a *Wenz* or *Solo*. The amounts are higher when players win with a large margin.<sup>8</sup> Finally, as mentioned earlier, each knock on the table after the distribution of the first four cards and each “Contra” results in a doubling of payouts. The cents that players accumulate over the rounds constitute their score.

After each round, each player decides if they want to stay in for another game or leave the table. Distribution of new cards only starts when four players sit at the table.

## 1.2 The online platform

Our data was provided by *sauspiel.de*, the largest online *Schafkopf* gaming community. The platform was founded in 2007 by a group of four students with the goal of bringing this tra-

---

<sup>4</sup>*Sauspiel* thus involves a competition between teams, while *Solo* and *Wenz* require the individual to compete alone against all other players at the table. Consequently, the stakes are also higher in the latter two types. The difference between *Solo* and *Wenz* is that only Jacks count as trumps for a *Wenz* game while the *Solo* is more similar to the *Sauspiel* in that Queens, Jacks and one designated color count as trumps. The partner of a *Sauspiel*’s initiator cannot be freely chosen but is randomly determined by the initiator calling a specific suit of ace. Team composition in *Sauspiel* games is thus only revealed when the specific ace is played out. Team composition for the more competitive *Wenz* or *Solo* game is immediately revealed as the initiator plays against the three other players.

<sup>5</sup>Players have to contribute a card of the same color as the first played card of the trick, but can play any other card if they don’t possess a card of the same color. If the first played card of the trick is a trump card, the other players also have to play a trump card if they still possess one.

<sup>6</sup>If a trick contains multiple Queens or Jacks, their rank order is determined by suit.

<sup>7</sup>11 for the Ace, 10 for the 10, 4 for the King, 3 for the Queen, 2 for the Jack, and 0 for all other cards.

<sup>8</sup>The above amounts are increased by 10 cents if the losing party obtained less than 30 points and by 20 cents if no tricks were won by the losing party. Furthermore, if the winning party had a sequence of at least the three highest trumps, the sum is increased by 10 cents per trump.

ditional card game to the online world. The founders implemented the online version of the game with exactly the same rules as for regular *Schafkopf*, lowering entry costs for experienced players of the game. The platform became quickly popular and has already hosted more than 500 million games as of 2015.

First-time players have to register a profile on the platform, which includes choosing a user name and customizing a male or female avatar.<sup>9</sup> Registration and the use of the platform is free. Virtual ‘cents’ collected in the online games have no value outside the platform. As *Schafkopf* is always played by a group of four players, each player joins a virtual table of four. Figure A2 in the Appendix shows the table selection stage: players can either set up a new table by clicking on the plus symbol and wait until three other players join, or join a table with less than four players. As the tables usually fill up within a few seconds and because at this stage no information on other players’ past performance is displayed, there is little room for strategically selecting the set of players to share the table with.

Once four players have joined a table, the regular game begins: as described in the previous section, the players decide if they want to raise the stakes by knocking after the first four cards are distributed, then announce if they want to initiate a game after all eight cards have been dealt, and if at least one player announces a game, all eight tricks are played out consecutively. Players make their choice of initiating a game being informed of their own card quality, the decision of the other players at the table to raise stakes (which can serve as an indicator of their card quality), the past performance of other players (indicated by their cumulative scores displayed below their user names, as shown in Figure A1), and the gender composition at the table (shown by the avatars). The median game duration including the selection stage is only 81 seconds, so players usually stay at the same table for multiple game rounds. If one player leaves the table, the remaining players can stay at the same table and wait until a new participant joins them.

We observe all games played between September 5<sup>th</sup> 2007 and January 9<sup>th</sup> 2008. For each round during these five months, we observe the type of game played (*Sauspiel*, *Wenz* or *Solo* or no game if everyone “passed”), who initiated the game and who has been in a partner (for *Sauspiel*) or opponent role. We can also see who raised the stakes of the game by “knocking” and in which order players announced their game choice (i.e. players’ position at the table). For each initiator we observe whether they were given “Contra” by their opponent(s), but the identity of the exact opponent(s) who gave the “Contra” is not recorded. Therefore the individual decision to give “Contra” can not be analyzed in our setting.

If a game is played, our dataset also contains the exact cards played out by each player, allowing us to assess whether good or bad cards were dealt to the players. We calculate a joint measure of card quality by regressing multiple indicators of card value<sup>10</sup> on the probability of winning

---

<sup>9</sup>Players can voluntarily also register their gender, date of birth and their ZIP code. In our sample, 22% of all players provided this information.

<sup>10</sup>The indicators are the cumulative point value of all owned cards as well as dummy variables for the number of trump cards, the number of suits a player does not possess (as this allows to go in with a trump if another player plays this suit), the number of aces, and the consecutive number of highest trumps in a row.

a *Sauspiel* and then using the coefficients to derive the joint measure of card quality.<sup>11</sup> If all players decide to “pass” and cards are therefore not played out, the distribution of cards is not recorded.

For each game, we observe the winner(s) and the number of points achieved (between 0 and 120) as well as the final score in cents for each participant. As we have seen in Figure A1, each player’s cumulative score is publicly displayed on the screen during the games and thus influences the status in the community. The displayed score provides an imperfect indication of a player’s skills and performance because these scores can be reset to zero whenever they are falling below zero. Our dataset does not contain these displayed scores. However, we do observe the actual scores and can use them to infer the displayed scores of all players.

### 1.3 Data

This section introduces our dataset and the most important variables in more detail. Table 1 provides an overview of our data. The dataset contains over 4 million games recorded from the perspective of each player at the table, resulting in 16,655,344 observations in total. Our data is generated by more than 15 thousand individual players. The share of female players<sup>12</sup> in the community is low, around 8.5%, reflecting *Schafkopf* being a traditionally “masculine” activity. Female players are, however, very active on the platform: since the number of games per player is much higher for women (mean: 1,745) than for men (mean: 1,033), they generate 13.52% of all observations (see also Figure A3 for the distribution of the number of games per player).

As mentioned earlier, our dataset does not contain much information about the characteristics or background of the players. For a subsample of 3,323 registered users we observe their age and the ZIP code of their residence. Based on this subsample we do not observe a substantial difference between men and women in terms of their age: as shown in Panel B of Table 1, the mean age is 30.19 for men and 29.80 for women. Figure A4 in the Appendix, depicting the age distribution of our players, shows that our sample is very diverse, with some players younger than 15 and others older than 70 years. Only a minority of the registered users come from Munich, and women are slightly more likely than men to live in the state capital.

Table 2 tabulates the gender composition of tables for all rounds. In more than half of the cases there were only male players at the table, approx. 36% of games involved one female player, and less than 0.03% of the rounds were played by four women at the table (this latter share, however, still corresponds to 1,110 games).

Even though we do not observe the employment or marital status of the players in our sample, we can still draw some cautious inferences regarding their lifestyles based on the hour of the day when they play games on the online platform. Figure A5 in the Appendix suggests that there is no systematic difference between men and women in the game community: both are

---

<sup>11</sup>The results are stable to including the different measures of card quality separately instead of the joint card quality measures in the regressions performed later.

<sup>12</sup>In our analysis we talk about female and male players when discussing users who registered a female or male avatar. We discuss in more detail to what extent this proxy provides an accurate representation of players’ true gender in Section 1.4.

most likely to play in the evening hour. Moreover, we find no indication for women playing more often during the typical working hours, suggesting that they are as likely as men to be employed.

Table 1: OVERVIEW OF THE DATASET

<i>Panel A</i>					
	MALE	%	FEMALE	%	TOTAL
Number of observations	14,402,768	86.48	2,252,576	13.52	16,655,344
Number of players	13,941	91.52	1,291	8.48	15,232
<i>Panel B</i>					
	MALE		FEMALE		
	Mean	Median	Mean	Median	
Number of games per player	1,033	238	1,745	447	
Age	30.19	28	29.80	27	
Location Munich	0.113	-	0.178	-	

Table 2: GENDER COMPOSITION OF GAMES

Women at table	N	%
0	2,308,689	55.45
1	1,491,212	35.81
2	333,553	7.96
3	31,273	0.75
4	1,110	0.03
Total	4,163,837	100

#### 1.4 Evaluation of our setting

Both the large dataset at hand and the particular features of the *Schafkopf* game are important ingredients for our contribution to the literature. A major advantage of our context is that we study naturally occurring data. The experimenter demand effect is absent and participants’ choices are based on a good understanding of the rules of the game they find very natural. The large dataset, both in terms of the number of individual players and the number of observations per player, allows us to detect small effect sizes and heterogeneous effects. It also enables the convincing measurement of potential null-effects. This kind of precise measurement would be difficult to obtain in a typical laboratory experiment with a limited number of participants. Moreover, unlike most lab experiments that are based on student samples, our dataset covers a wider age range and most likely also different levels of education. People self-select in the sample and, due to the character of the game, this spontaneously results in a male-dominated group (which is a good approximation for the environment at the top in corporate hierarchies in most labor markets). The features that participants earn only “virtual” cents and that players can be identified by their user names (which enables reputation building) ensure that players are very highly intrinsically motivated and care about their ranks.

The *Schafkopf* game has certain appealing features for studying gender differences in competition and risk taking and their likely consequences for performance and success. In our setting,

players face real interaction: the actions taken by a player’s opponent and partner actually affect the player’s chances of winning. Success in this set-up thus depends on strategic uncertainty (the behavior of other players at the table), non-strategic uncertainty (the distribution of card quality between players) and individual ability (game-specific skills). This feature is missing from related laboratory studies where tournament winners are typically selected based on their performance in an individually executed task.<sup>13</sup>

Moreover, in our data the interaction between players is repeated hundreds, often thousands of times, and players receive feedback after each round. At the end of each game they learn whether or not they have won and by how much their scores increased/decreased as a consequence. As mentioned before, players’ ranks in terms of displayed cumulative scores are common knowledge. Therefore, players do not only receive updates on their own absolute scores, but also on their position compared to other players. Feedback on one’s performance both in absolute and relative terms is a characteristic feature of labor and other markets. It is also a relevant feature of the data when studying gender differences in competitiveness because there is evidence that feedback induces more competitive choices among high-ability women (Wozniak et al., 2014). Despite all these appealing features of the collected data, we also have to deal with a number of limitations. First, little is known about the characteristics of our subjects (as discussed before, self-reported information on age and residence is available, but only for a sub-sample of the players). Our measures would be biased if female players had systematically different unobserved characteristics than males and these characteristics were correlated with the preferences we study.

Second, the gender measure we use (the gender of the avatar chosen by the players) is self-reported and could potentially be misreported. For a subsample of 3,323 individuals we also observe the gender they administer at registration and can compare this with the gender of their avatar. In 97 percent of the cases the two are identical.<sup>14</sup> Based on this information we assume that most peoples’ avatars are a truthful representation of their sex. Nevertheless it might be theoretically possible that people misrepresent their sex both at registration and when choosing their avatar. This would increase the overall percentage of misrepresented avatars somewhat. It might also be that people push the wrong button when registering, while their avatar (that can be updated all the time) is the correct representation of their gender. That would decrease the share of misrepresented avatars in the total sample. In any case, misrepresentation seems quite low. If it occurred at random, its only consequence would be less precise estimates. However, our estimates would be biased upwards if the most competitive female players systematically played with a male avatar and the least competitive male players played with a female avatar. Anecdotal evidence from *Schafkopf* players suggests that such gender misrepresentation is rare,

---

<sup>13</sup>In our setting, players do not make a choice between “tournament” and “piece rate” as common in lab experiments. In *Schafkopf*, the alternative of not initiating a game is either participating as partner or opponent in a game initiated by someone else, or playing no game at all. We find that this feature brings our context closer to real-life situations where a “safe option” or a non-competitive track is often unavailable. It also allows us to assess the performance of the less competitive players under tournament incentives.

<sup>14</sup>Approximately two-thirds of the “mismatches” result from players registered as female choosing a male avatar.

but this claim is admittedly untestable given our data.

A third, more minor, disadvantage of our setting is that the game is too complex for us to derive the “optimal” strategy and behavior in each situation. Therefore we can not conclude whether women initiate games “too little” and men “too much”. Instead, we can infer the *relative profitability* of their playing strategies by comparing the scores collected by men and women (controlling for their playing ability).

## 2 Results

In this section we test whether female players in our dataset behave differently than male players when it comes to risk taking and competition. In particular, we focus on two decisions players make in the selection stage of the game: whether to increase the stakes of the game by “knocking”, and whether to actively initiate a game (and which type). We then consider whether there are skill differences in the playing stage between men and women in our sample, where ‘skill’ refers to players’ ability to win games and collect points *given their role and the game type*. We also compare the success of male and female playing strategies by analyzing the scores players accumulate over the games. Finally, we speculate whether women’s differential selection choices may be attributed to gender differences in risk aversion and competitiveness. Throughout the analysis, we use the following terminology: “focal player” refers to the individual whose perspective we consider in the analysis (remember that all games are recorded from the viewpoint of all four participants), and “opponents” denote the three other players sitting at the table. We decided to use the general term “opponent” because even though in the *Sauspiel* game one of the fellow players will actually be the initiator’s partner, in the selection stage it is not yet known who this partner may be.

### 2.1 Descriptive evidence

Table 3 provides simple descriptive evidence for gender differences in playing behavior. It shows a large and significant gender gap in the propensity to raise the stakes: while female players only “knock” in 20.52% of the rounds they play, men do so in 28.11% of the cases. Women also play significantly less often in the initiator role than men: the difference is 3.34 percentage points (this amounts to a gender gap of approximately 22%, given that women play as initiators 15% of the rounds). Consequently, women are more likely than men to play as opponents or as partners or not to play any game in the given round.

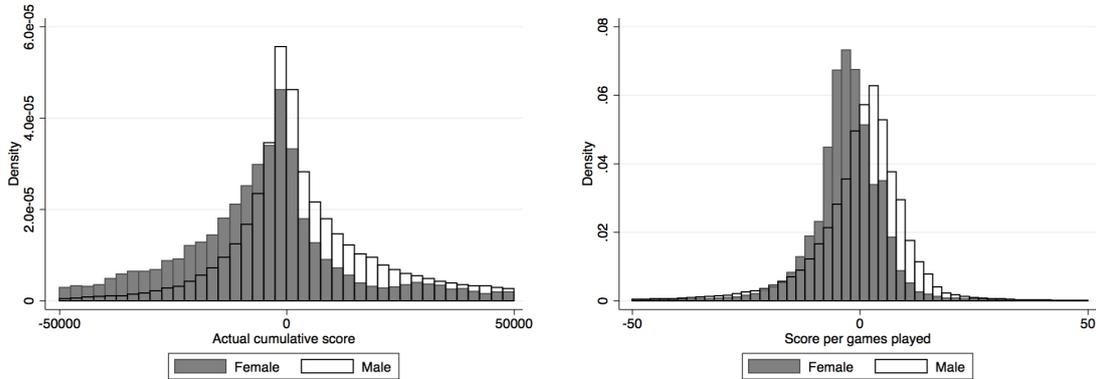
Table 3 also shows winning probabilities for players in each role. Female players are somewhat more likely than males to win games in the initiator and partner role (the difference is 2.02 and 0.37 percentage points, respectively), and underperform slightly (by 0.4 percentage points) compared to men in the opponent role.

Table 3: DESCRIPTIVE EVIDENCE: Gender differences

	MALE	FEMALE	Difference
Raise stakes (“knock”), %	28.11	20.52	-7.59***
Role in game, %			
Initiator	18.37	15.03	-3.34***
Partner	12.14	12.61	0.47***
Opponent	41.44	42.22	0.78***
None	28.05	30.14	2.09***
Games won in role, %			
Initiator	76.39	78.41	2.02***
Partner	79.82	80.19	0.37***
Opponent	24.39	23.99	-0.40***
Score (mean)			
Cumulative	12,692	-7,504	-20,196***
Displayed	20,021	12,142	-7,879***

Notes: Significance of differences from t-tests with unequal variances; \*\*\* p<0.001, \*\* p<0.01, \* p<0.05

Finally, Table 3 shows that despite playing more games and having similar winning probabilities, women on average accumulate substantially lower scores (both displayed and actual) than men in our sample: the mean score of male players is 12,692 while it is -7,504 for women. Figure 1a displays the distribution of actual cumulative scores by gender, while 1b depicts the scores collected per games played. Both figures confirm that female players are on average less successful in the community than males.



(a) Cumulative scores (truncated at  $\pm 50,000$ ) (b) Scores per games played (truncated at  $\pm 50$ )

Figure 1: Distribution of cumulative scores, by gender

## 2.2 Behavior in the selection stage

We continue by analyzing in more detail the finding that men and women play differently in the selection stage. Table 4 reports marginal effects<sup>15</sup> calculated from probit models analyzing the propensity to initiate a game (columns (1)-(3)) and to raise the stakes by “knocking” (column (4)).<sup>16</sup> The table also contains the “baseline” initiation/knocking rates, i.e. the predicted

<sup>15</sup>Estimated coefficients are displayed in Table A1 in the Appendix.

<sup>16</sup>To account for the fact that observations in our sample may be correlated in two non-nested dimensions (we observe multiple observations for the same individual, and four people play the game and interact at a single

probabilities for male players, evaluating all coefficients (other than gender) at their mean values. Column (1) displays results from a specification that only includes the focal player’s position at the tables as a control variable. We find that there is indeed a significant and sizable gender gap in the willingness to initiate games: while male players on average have an 18.3% likelihood of starting a game, being female is associated with a 3.3 percentage point drop in initiation rates.

Table 4: LIKELIHOOD OF INITIATING A GAME AND RAISING STAKES

VARIABLES	Initiate a game			Increase stakes
	(1)	(2)	(3)	(4)
Female player	-0.033*** (0.003)	-0.034*** (0.003)	-0.029*** (0.003)	-0.068*** (0.008)
Position at table = 2	-0.022*** (0.000)	-0.021*** (0.000)	-0.021*** (0.000)	-0.010*** (0.001)
Position at table = 3	-0.031*** (0.000)	-0.030*** (0.000)	-0.030*** (0.000)	-0.010*** (0.001)
Position at table = 4	-0.035*** (0.001)	-0.035*** (0.001)	-0.035*** (0.001)	-0.008*** (0.000)
Female opponent		-0.003*** (0.000)	-0.004*** (0.000)	-0.006*** (0.001)
Num. opponents knocked = 1		-0.110*** (0.000)	-0.109*** (0.000)	
Num. opponents knocked = 2		-0.170*** (0.001)	-0.169*** (0.001)	
Num. opponents knocked = 3		-0.202*** (0.001)	-0.201*** (0.001)	
Num. games played (log)			-0.005*** (0.001)	0.003 (0.001)
Rank in score = 2			-0.013*** (0.001)	-0.029*** (0.003)
Rank in score = 3			-0.020*** (0.002)	-0.052*** (0.004)
Rank in score = 4			-0.024*** (0.002)	-0.081*** (0.005)
Predicted prob. male players	0.183*** (0.001)	0.176*** (0.001)	0.175*** (0.001)	0.279*** (0.002)
Observations	16,655,344	16,655,344	16,655,344	16,655,344
Number of players	15232	15232	15232	15232
Pseudo $R^2$	0.00230	0.0312	0.0317	0.00748

Notes: The table displays marginal effects at the means from probit models (for factor variables, calculated as the discrete change from the base level). Dependent variable in columns (1) - (3): initiate any game type; column (4): increase the stakes of the game by “knocking”. As a baseline, we present predicted probabilities for male players (evaluating all other variables at their means). Omitted categories for covariates: position at table = 1; female opponent = 0; rank in score = 1; num. opponents knocked = 0. Standard errors are clustered on player ID and are reported in parentheses. \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

table), we cluster standard errors at the individual level, and add table level controls (see Cameron and Miller (2015)).

Column (2) and (3) show that this gender gap is remarkably robust to the inclusion of covariates. Controlling for the presence of females among the opponents, the number of opponents who “knocked” in the selection stage, the focal player’s relative rank in terms of displayed cumulative scores or the focal player’s experience measured by the (natural logarithm of) the number of games he/she has played before, the estimated gender difference in the likelihood to initiate a game remains in the range of 2.9–3.4 percentage points. As we see in column (4), the gender gap in knocking behavior is even larger: all else equal, female players are 6.8 percentage points less likely to increase the stakes, compared to a “baseline” predicted probability of 27.9% for men.

As noted before, cards dealt to players in rounds where no-one initiated a game are not recorded in our dataset. It is therefore not possible to include card quality as a control variable in the above regression explaining initiation decisions. We can, however, exploit the very large number of observations in our dataset and compare the theoretical card quality distribution that occurs when cards are dealt at random with the actual card quality distribution observed for initiated games. This allows us to make inferences about the way card quality influences the entry decision on average in our sample. Figure 2 shows that the gender gap in game initiation rates persists throughout the card quality distribution, including the best cards as well.<sup>17</sup>

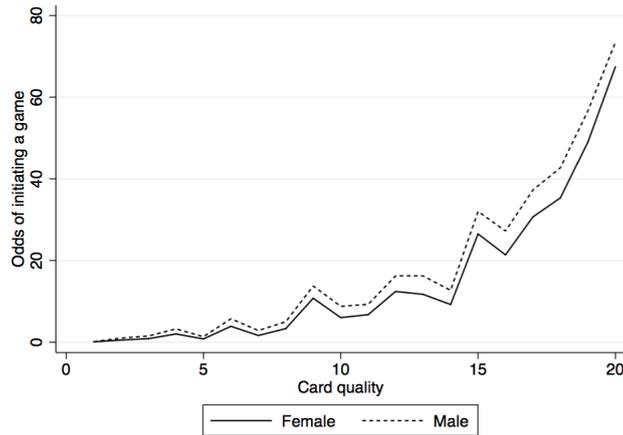


Figure 2: Odds of initiating a game by card quality vigintiles, separately by gender

We consider how the gender gap in initiation rates differs by game type. As discussed in Section 1.1, players can choose between three game types: *Sauspiel* involves a team competition, whereas in *Wenz* and *Solo* games the initiator competes alone against all three other players at the table. In the latter two game types the stakes are also higher. We conjecture that the gender gap in initiation rates should be higher in the more competitive and risky *Wenz* and *Solo* games than

<sup>17</sup>The x-axis of the graph in Figure 2 shows card quality vigintiles. The apparent non-monotonicity of initiation rates by card quality results from the way we created the proxy for card quality. The card quality measure is calculated from several proxies for card quality, but is likely still missing some information that is related to the choice of initiating a game. So while the odds of initiating a game generally increase with card quality, uncaptured differences in card quality that are correlated with the captured measures can lead to the observed non-monotonic relationship.

in *Sauspiel*. To test our hypothesis, we estimate a multinomial logit model explaining the choice to initiate one of the three game types (using the category *initiate no game* as our baseline). Table 5 displays marginal effects estimated from this model. We find that in case of *Sauspiel*, male players’ predicted probability of initiating this game type is 12.4%, and the gender gap is 1.4 percentage points. In case of *Solo* games, the gender gap is 0.9 percentage points - a large difference given that men on average initiate this game type only 3.6% of the rounds. Similarly, a gender gap of 0.6 percentage points in *Wenz* initiation rates amounts to a sizable disparity compared to men’s “baseline” initiation rate of 2.3%.

Table 5: GENDER GAP IN INITIATION RATES, BY GAME TYPES

VARIABLES	Sauspiel (1)	Wenz (2)	Solo (3)
Female player	-0.014*** (0.002)	-0.006*** (0.001)	-0.009*** (0.001)
Predicted prob. male players	0.124*** (0.001)	0.023*** (0.000)	0.036*** (0.000)
Covariates included		✓	
Observations		16,655,344	
Number of players		15232	
Pseudo $R^2$		0.00220	

Notes: The table displays marginal effects from a multinomial logit model. The dependent variable is *game type initiated* with four categories: Sauspiel; Wenz; Solo; none (baseline). Covariates included (unreported): position at table, female opponent, experience (log. number of games played), rank in score, num. opponents knocked. Omitted categories for covariates: position at table = 1; female opponent = 0; rank in score = 1; num. opponents knocked = 0. Standard errors are clustered on player ID and are reported in parentheses. \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

We also test whether the overall gender gap in initiation rates results from female players responding differently to the characteristics or behavior of their opponents. Table A2 in the Appendix, showing estimated coefficients<sup>18</sup> from probit models with gender interaction terms, suggests that this is not the case. We find that neither men nor women respond strongly to the presence of women at the table (having at least one female among the three opponents increases the initiation rates very slightly, and the gender interaction is insignificant).<sup>19</sup> Players are influenced by their relative rank at the table: both men and women are less likely to initiate games when they are ranked worse than their opponents, and women are especially deterred from initiating when they have the lowest accumulated score at the table. On the other hand, there is no gender difference in the response to opponents’ knocking behavior: male and female players are equally discouraged from initiating a game when their opponents raise the stakes. Finally, Figure 3 demonstrates that experience affects the behavior of men and women differently. Initiation rates decrease for women with the number of games they have played, while

<sup>18</sup>We report estimated coefficients - instead of marginal effects - because obtaining and interpreting marginal effects is less straightforward in case of models including interaction terms of factor variables. The coefficients in this table may be compared to the estimates reported in column (3) of Table A1.

<sup>19</sup>We also test whether we can replicate the finding that women are more competitive in a single-sex environment (Gneezy et al., 2003) by studying the initiation rates of women at female-only tables. We find no evidence in our data for female players being more likely to initiate a game in the absence of men.

the relationship is U-shaped for men. As a result, the gender gap is largest among the most experienced players: practice and feedback, instead of closing the gap, actually widens it. This divergent pattern by experience could tentatively be explained by gender differences in response to losing a competition (Buser, ming): if losses induce men but not women to subsequently take more challenges, the gender gap in game initiation rates should widen over time.<sup>20</sup>

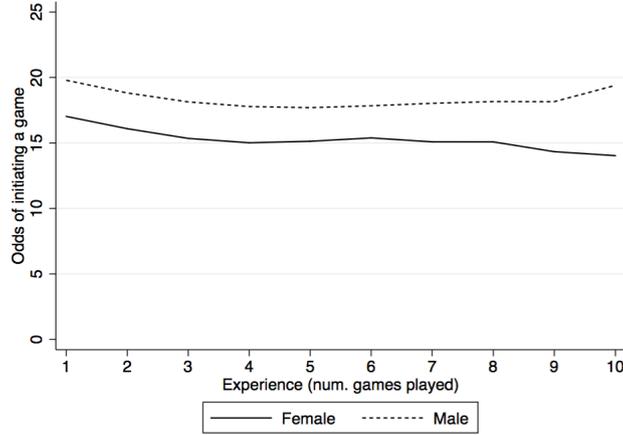


Figure 3: Odds of initiating a game by experience deciles, separately by gender

### 2.3 Performance in the playing stage

Having considered players’ behavior in the selection stage, we continue by analyzing their performance in the playing stage. We study whether the large score difference between men and women in our sample can be attributed to differences in playing skills. We start by focusing on winning probabilities of players in games they initiated. The first column of Table 6 repeats the result from Section 2.1 in a regression framework: women are slightly but significantly more likely to win games they started themselves. While the predicted probability of winning for male players is 76.4%, being female is associated with a 2.0 percentage points increase in winning odds. When we include covariates such as position at the table, the presence of a female opponent or the number of rounds played on the platform before, the gender difference is reduced to 1.4 percentage points (column (2)). Controlling for the initiator’s and opponents’ quality of cards we find that male and female players are equally likely to win: the estimated coefficient for female player is reduced to zero and loses significance. The combined results of this table suggest that women have on average better cards than men when they initiate games. More importantly, column (3) shows that given card quality, there is no gender difference in the ability to win self-initiated games.<sup>21</sup> This finding is illustrated in Figure 4.

<sup>20</sup>To correct for the fact that a large share of our observations comes from a minority of players who play thousands of games in the period we observe (thus the most active and experienced players have a disproportionately large influence on our results), we estimate a model explaining initiation behavior using weights that are inversely proportional to the total number of games played by each individual. The gender gap remains significant and similar in size even after this correction.

<sup>21</sup>This result is unchanged when we correct for the fact that players in our dataset differ in the total number of rounds they played. Estimating a probit model with weights that are inversely proportional to the total number

Table 6: LIKELIHOOD OF WINNING SELF-INITIATED GAMES

VARIABLES	No controls (1)	With controls (2)	Card quality (3)
Female player	0.020*** (0.003)	0.014*** (0.003)	0.001 (0.003)
Position at table = 2		-0.008*** (0.001)	-0.021*** (0.001)
Position at table = 3		0.014*** (0.001)	-0.003*** (0.001)
Position at table = 4		0.034*** (0.001)	0.015*** (0.001)
Num. games played (log)		0.015*** (0.001)	0.015*** (0.000)
Female opponent		0.002*** (0.001)	-0.003*** (0.001)
Card quality			0.151*** (0.003)
Opponents' card quality			-1.073*** (0.005)
Partner's card quality			0.250*** (0.003)
Contra			-0.239*** (0.001)
Predicted prob. male players	0.764*** (0.001)	0.766*** (0.001)	0.799*** (0.001)
Observations	2,984,123	2,984,123	2,984,123
Number of players	14302	14302	14302
Pseudo $R^2$	0.000214	0.00473	0.136

Notes: The table displays marginal effects at the means from probit models (for factor variables, calculated as the discrete change from the base level). Dependent variable: odds of winning self-initiated games. As a baseline, we present predicted probabilities for male players (evaluating all other variables at their means). Omitted categories for covariates: position at table = 1. For the variable partner's card quality, missing values (in case of Solo and Wenz games) are imputed with zeros, and an additional indicator variable for missing values is included. Standard errors are clustered on player ID and are reported in parentheses. \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

of games played, the difference between the winning odds of male and female players is small and only marginally significant.

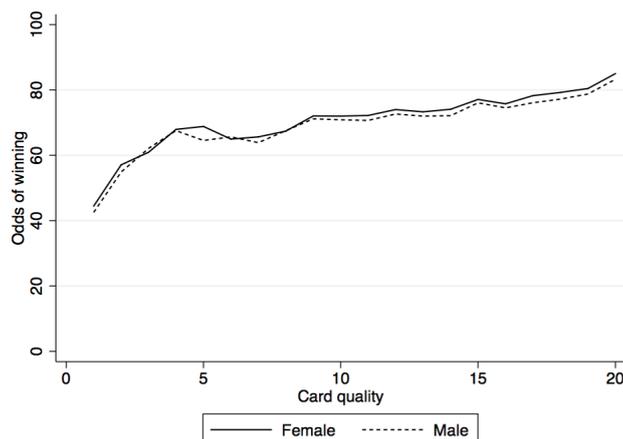


Figure 4: Odds of winning a game (conditional on initiating it) by card quality vigintiles, separately by gender

Besides winning a game, we can also analyze points collected in a given round. Column (1) of Table 7 confirms that women are as successful players as men according to this continuous measure of performance: female players collect as many points per initiated game as men. Studying self-initiated games we have thus found that female players perform as well as men. This result, however, suffers from potential endogeneity problems. Players self-select into the initiator role and this selection is unlikely to be random: the unobservable characteristics that influence initiating behavior could also affect winning probabilities, and could be correlated with gender. We therefore also consider performance in a situation where the scope for self-selection is virtually absent: players in the fourth position at the table playing as opponents in *Solo* games. To see why the opponent role is exogenously assigned in these roles, consider the following. Players in the fourth position are the last to announce their intent to initiate games. Therefore the only way for them to prevent being an opponent in a game started by someone before them at the table is to call a higher game type than the one already initiated. Since *Solo* is the highest type, there is no way for the last player at the table to avoid the opponent role in this game type. Column (2) of Table 7 analyzes the winning odds of opponents in the fourth position of *Solo* games and confirms that there is no significant gender difference in performance in these exogenously assigned roles, either.<sup>22</sup> We can thus conclude that given their cards and their role in the game, male and female players in our sample perform equally well: there is no gender gap in “on-task” performance.

<sup>22</sup>Another way to address the endogeneity of winning probabilities in self-initiated games would be to estimate a Heckman sample selection model. However, we did not find a variable in our dataset that would influence game initiation but not performance in the game. In the absence of such an exclusion restriction, identification of the estimated parameters depends on untestable functional form assumptions. Moreover, the most important predictor of winning, card quality, is only observed in cases where a game is actually initiated. Adding it as a control variable in the main but not in the selection equation would potentially make it an endogenous covariate. Based on these considerations we decided not to include a probit model with sample selection in our analysis.

Table 7: DIFFERENT PERFORMANCE MEASURES

VARIABLES	<i>Full sample</i>	<i>Players in 4<sup>th</sup> position</i>
	Points in initiator role (1)	Winning Solo in opponent role (2)
Female player	0.176 (0.158)	0.001 (0.002)
Position at table = 2	-1.516*** (0.035)	
Position at table = 3	-0.440*** (0.035)	
Position at table = 4	0.587*** (0.035)	
Num. games played (log)	0.927*** (0.033)	0.001 (0.001)
Female opponent = 1	-0.297*** (0.032)	-0.004** (0.002)
Card quality	19.901*** (0.168)	0.074*** (0.004)
Opponents' card qual.	-58.542*** (0.262)	-0.636*** (0.005)
Partner's card quality	27.250*** (0.148)	
Contra	-17.080*** (0.081)	0.350*** (0.003)
Constant	117.183*** (0.296)	
Predicted prob. male players		0.290*** (0.001)
Observations	2,984,123	427,093
Number of players	14302	12086
Adjusted/Pseudo $R^2$	0.258	0.0863

Notes: The table displays estimated coefficients from an OLS (column (1)) and marginal effects at the means (for factor variables, calculated as the discrete change from the base level) from a probit model (column (2)). Column (1): dependent variable: points collected in a self-initiated game; covariates: for the variable partner's card quality, missing values (in case of Solo and Wenz games) are imputed with zeros, and an additional indicator variable for missing values is included. Column (2): estimated on the subsample of players in the fourth position at the table. Dependent variable: odds of winning a Solo game for players in the opponent role. As a baseline, we present predicted probabilities for male players (evaluating all other variables at their means). Omitted categories for covariates: position at table = 1. Standard errors are clustered on player ID and are reported in parentheses. \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

## 2.4 Explaining the gender difference in scores

In the previous two subsections we have established that female players indeed “shy away”: they make less competitive and risky choices in the selection stage. Furthermore, we have shown that there is no gender difference in winning probabilities in the playing stage once we control for card quality and roles. This latter result is puzzling given the difference in cumulative scores between male and female players (see Figure 1). In the following we explain two ways in which

gender differences in choices lead to lower scores for women despite no gender difference in actual winning ability.

First, as a result of their lower propensity to initiate a game, female players end up more often in the difficult opponent role and less often in the easier initiator role. Consequently, even though the winning probabilities are the same for male and female players *in each role*, the *overall* winning rates of women are lower.

Moreover, scores collected in the game are calculated as the sum of the cents earned, and women are less effective in turning the points they win into cents. Note that the conversion between points and cents depends on the game type (*Wenz* and *Solo* pay more) and the chosen stakes of the game (determined by the number of ‘knocks’ and ‘Contras’). Since women disproportionately shy away from the more competitive (and more lucrative) game types and fail to increase the stakes even when their cards are good, they profit less from winning than men do.

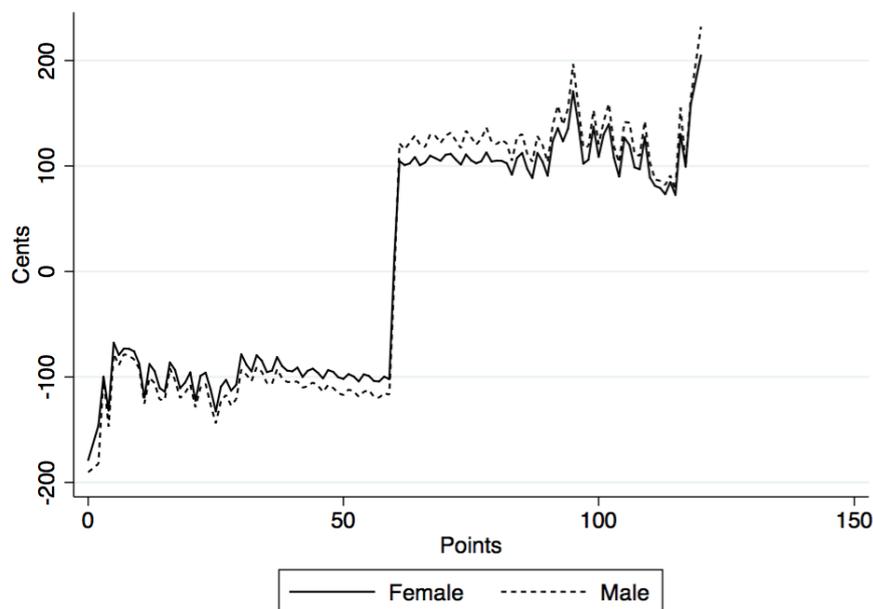


Figure 5: Conversion between points and cents, separately by gender

Figure 5, plotting the mean cents earned against the points collected, illustrates the above reasoning. It shows that female players gain less from winning and lose less from losing a game than males (note that collecting at least 61 points out of 120 ensures winning). The difference is smaller in absolute terms in case of winning than in case of losing: as winners, men earn on average 19.5 cents more than women for the same number of points, while as losers women lose 14.3 cents less than men, given equal points. Over the course of several games, the difference accumulates (exacerbated by the above-discussed result that women have lower overall winning rates).

## 2.5 Drivers of the gender gap in choices

In our analysis we have shown that female players have a different playing style than men: they have a lower propensity to initiate a game and to raise stakes. We now consider whether these differences can be attributed to gender differences in preferences towards risk and competition, or they are rather driven by alternative forces.

We first discuss whether “shying away” is a result of discrimination against female players. Table 8 suggests that this is unlikely to be the case: men do not play more aggressively against female opponents. To the contrary, male players are actually slightly *less* likely to initiate the most competitive game type (*Solo*) (column (1)) and are *less* inclined to raise the stakes (column (2)) when there is a woman at the table (both differences are negligible in size). In column (3) we consider players who have an all-male group of opponents and show that female players are less likely to receive “Contra”. We thus find no sign of women being ‘penalized’ for participating in this traditionally masculine game. Additionally, we find (in unreported analysis) the gender gap in initiation rates to be the largest among players in the first position (i.e. who are the first to announce their intent to play) where it is certainly not a response to the actions of others at the table. Taken together, these results imply that women’s different choices are not a response to negative discrimination on the side of male players.

A gender gap in confidence is another candidate explanation for the differences observed between male and female playing styles (Kamas and Preston, 2012). As shown in Table A2, women indeed react more to negative feedback on their relative performance than men do (they are particularly deterred from initiating when they have the worst rank at the table). Controlling for the interaction term between female and rank (column (3)), the gender gap in game initiation rates reduces slightly but remains significant, suggesting that confidence alone cannot account for the differences between male and female behavior in the selection stage.

It could also be argued that female players are simply less skilled than males in identifying rounds when it is beneficial to initiate a game or to increase the stakes. While we can not rule out this explanation, observing the initiation rates by card quality in Figure 2 suggests that this is not the main driver. The skill of identifying opportunities should play the most important role in game initiation decisions in case of intermediate quality cards: even the least able player is unlikely to start a game with obviously bad cards and will probably not fail to initiate when he/she is dealt the best possible cards. The gender gap in initiation rates, however, is not the largest in case of medium quality cards (in absolute terms, it is largest for the highest, in relative terms, for the lowest quality cards).

Finally, women’s less offensive playing style could be driven by their relative disinterest in winning. Women may initiate and raise the stakes less often because they get less utility from high scores than men do. This reasoning holds especially since the game is played for “virtual cents”, not for real money. However, an (unreported) analysis of hazard rates suggests that women are even more likely to leave a table after a loss than men. This particularly strong negative response to losing a game suggests that women care at least as much about their performance in the game as male players.

Table 8: REACTION TO FEMALE PLAYERS

VARIABLES	<i>Male players</i>		<i>Players with all-male opponents</i>
	Initiate Solo (1)	Raise stakes (2)	Receive Contra (3)
Female player			-0.007*** (0.001)
Female opponent	-0.002*** (0.000)	-0.013*** (0.001)	
Position at table = 2	-0.004*** (0.000)	-0.009*** (0.001)	-0.000 (0.000)
Position at table = 3	-0.004*** (0.000)	-0.010*** (0.001)	-0.000 (0.000)
Position at table = 4	-0.001*** (0.000)	-0.008*** (0.001)	-0.000 (0.000)
Num. games played (log)	0.000* (0.000)	0.013*** (0.001)	-0.002*** (0.000)
Num. opponents knocked = 1	-0.025*** (0.000)	-0.118*** (0.001)	0.012*** (0.000)
Num. opponents knocked = 2	-0.045*** (0.000)	-0.219*** (0.001)	0.024*** (0.000)
Num. opponents knocked = 3	-0.060*** (0.001)	-0.281*** (0.003)	0.033*** (0.001)
Predicted prob. male players	0.033*** (0.000)	0.276*** (0.002)	0.040*** (0.000)
Observations	14,402,768	14,402,768	10,725,967
Number of players	13941	13941	14618
Pseudo $R^2$	0.0258	0.0258	0.00720

Notes: The table displays marginal effects at the means from probit models (for factor variables, calculated as the discrete change from the base level). Dependent variables: column (1): initiate Solo; column (2): increase the stakes of the game by “knocking”, column (3): receive “Contra”. Columns (1) and (2): subsample of male players; column (3): subsample of players with all-male opponents. As a baseline, we present predicted probabilities for male players (evaluating all other variables at their means). Omitted categories for covariates: position at table = 1; num. opponents knocked = 0. Standard errors are clustered on player ID and are reported in parentheses. \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

To conclude, we believe none of the above alternative explanations can account for the observed gender gap in playing styles. This leads us to conjecture that women’s lower game initiation and ‘knocking’ tendencies are to a large extent driven by gender differences in risk appetite and competitiveness. It is important to note, however, that women who have self-selected into the card game community are not averse to being *exposed* to risk and competition, but they prefer not to *initiate* such situations. Female players who do not start games themselves still enter the tournaments, only in the role of partner or opponent. Moreover, even though women are much less likely than men to increase the stakes, there is no difference in their response upon observing that their opponents ‘knock’ (see column (4) in Table A2). We find this phenomenon to be in line with women’s shame aversion (Ludwig and Thoma, 2012) or adherence to social norms (Gneezy et al., 2009).

### 3 Conclusions

Our results are consistent with the idea that even those women who have deliberately entered an uncertain tournament environment will subsequently “shy away” from actively seeking competition and risk. In the context we study, females are less inclined than males to initiate competitive games and to increase the stakes of the game. These findings replicate earlier results related to gender differences in preferences, albeit in a context and setup that is different from previous studies. Most notably, instead of a controlled laboratory experiment we obtain our results by studying a large set of naturally occurring data. The novelty of our study comes from showing that females’ lower perceived performance (i.e. lower cumulative displayed scores) are not due to differences in playing skills but to women’s less competitive and risky choices. Their more cautious playing behavior causes them to gain less and appear less competent in the game, despite the fact that given the games they play, female players do not underperform at all. We might - cautiously - extrapolate our findings to gender differences in behavior in real (labor) markets and thereby link gender differences in preferences to the slower advancement of women in organizational hierarchies.

Why is it problematic if women “shy away” and (therefore) have lower incomes and less steep career tracks in competitive labor markets? Perhaps women are willing to pay that price in order to avoid risk and competition. While this might explain the gender gap in choices, it is also possible that women are unaware of the *consequences* of their behavior in terms of forgone earnings and opportunities. For instance, Sandberg (2013) reports an example from Google where female engineers’ self-nomination rates rose to the same level as men’s after the management shared evidence with them about prior gender differences in nomination rates.

The consequences of gender differences in preferences are also important from a more general perspective. If qualified women refrain from using their full potential in terms of productivity by “shying away”, this could lead to efficiency losses. For instance, if the best candidate for a CEO position was a woman who was deterred by the cut-throat competition to reach the top, and therefore a lower quality male CEO was appointed instead, not only would the female candidate miss out on a valuable career opportunity but also the company could suffer from lower performance.

All in all, it is worthwhile to study and develop instruments that encourage women to take more risk and competition. As mentioned before, giving advice or sharing evidence of gender differences in behavior and their consequences might already be helpful. Encouragingly, a recent laboratory experiment shows that advice increases entry into tournaments among high-performing women (Brandts et al., 2015). It is important to note, however, that a lower appetite for risk and competition might actually be beneficial in certain contexts. Eckel and Füllbrunn (2015) find, for instance, that increasing the number of women as investors in an experimental financial market reduces overpricing. They support their results with a large meta-analysis of previous experiments that also shows a substantial negative correlation between the share of female traders in the market and the magnitude of price bubbles that occur. Such results suggest that besides (or instead of) coaching women to tolerate uncertainty and competition better, we could consider promotion mechanisms that rely less on self-initiated competition and risk taking.

## References

- Adams, R. B. and Funk, P. (2012). Beyond the glass ceiling: Does gender matter? *Management Science*, 58(2):219–235.
- Altonji, J. G. and Blank, R. M. (1999). Race and gender in the labor market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3, chapter 48, pages 3143–3259. Elsevier.
- Azmat, G. and Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics*, 30:32 – 40.
- Barsh, J. and Yee, L. (2012). *Unlocking the full potential of women at work*. McKinsey & Company.
- Bertrand, M. (2011). New Perspectives on Gender. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4, chapter 17, pages 1543–1590. Elsevier.
- Brandts, J., Groenert, V., and Rott, C. (2015). The impact of advice on women’s and men’s selection into competition. *Management Science*, 61(5):1018–1035.
- Buser, T. (forthcoming). The impact of losing in a competition on the willingness to seek further challenges. *Management Science*.
- Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *The Quarterly Journal of Economics*, 129(3):1409–1447.
- Cameron, C. A. and Miller, D. L. (2015). A practitioner’s guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.
- Card, D., Cardoso, A., and Kline, P. (2015). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. NBER Working Paper 21403.
- Carter, N. M. and Silva, C. (2010). Pipeline’s broken promise. In *The Promise of Future Leadership: A Research Program on Highly Talented Employees in the Pipeline*. Catalyst Org.
- Catalyst Org. (2015). Knowledge center: Women ceos of the sp500. <http://catalyst.org/knowledge/women-ceos-sp-500>. Accessed 30 January 2015.
- Charness, G. and Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83(1):50–58.
- Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–74.
- Dargnies, M.-P. (2012). Men too sometimes shy away from competition: The case of team competition. *Management Science*, 58(11):1982–2000.
- Delfgaauw, J., Dur, R., Sol, J., and Verbeke, W. (2013). Tournament incentives in the field: Gender differences in the workplace. *Journal of Labor Economics*, 31(2):305 – 326.
- Desvaux, G., Devillard-Hoellinger, S., and Meaney, M. C. (2008). A business case for women. *The McKinsey Quarterly*.

- Eckel, C. C. and Füllbrunn, S. C. (2015). Thar SHE Blows? Gender, Competition, and Bubbles in Experimental Asset Markets. *American Economic Review*, 105(2):906–20.
- Eckel, C. C. and Grossman, P. J. (2008). Differences in the Economic Decisions of Men and Women: Experimental Evidence. In Plott, C. R. and Smith, V. L., editors, *Handbook of Experimental Economics Results*, volume 1, chapter 57, pages 509–519. Elsevier.
- European Commission (2012). *She Figures. Gender in Research and Innovation: Statistics and Indicators*. Luxembourg: Office for Official Publications of the European Communities.
- Filippin, A. and Crosetto, P. (2014). A reconsideration of gender differences in risk attitudes. IZA Discussion Papers 8184, Institute for the Study of Labor (IZA).
- Flory, J. A., Leibbrandt, A., and List, J. A. (2014). Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions. *The Review of Economic Studies*, 82(1):122–155.
- Gneezy, U., Leonard, K. L., and List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, 77(5):1637–1664.
- Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, 118(3):1049–1074.
- Gneezy, U. and Rustichini, A. (2004). Gender and competition at a young age. *American Economic Review*, 94(2):377–381.
- Grosse, N. D., Riener, G., and Dertwinkel-Kalt, M. (2014). Explaining gender differences in competitiveness: Testing a theory on gender-task stereotypes. Working paper, University of Mannheim.
- Healy, A. and Pate, J. (2011). Can teams help to close the gender competition gap? *The Economic Journal*, 121(555):1192–1204.
- Institute for Leadership and Management (2011). Ambition and gender at work. IL-MAGW/0211.
- Ivanova-Stenzel, R. and Kübler, D. (2011). Gender differences in team work and team competition. *Journal of Economic Psychology*, 32(5):797–808.
- Kamas, L. and Preston, A. (2012). The importance of being confident: Gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization*, 83(1):82 – 97. Gender Differences in Risk Aversion and Competition.
- Kim, M.-K. and Polachek, S. W. (1994). Panel estimates of male-female earnings functions. *The Journal of Human Resources*, 29(2):pp. 406–428.
- Ludwig, S. and Thoma, C. (2012). Do Women Have More Shame than Men? An Experiment on Self-Assessment and the Shame of Overestimating Oneself. Discussion Papers in Economics 12905, University of Munich, Department of Economics.
- 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. (2011). Gender and competition. *Annual Review of Economics*, 3:601–630.

- Sandberg, S. (2013). *Lean In. Women, Work and the Will to Lead*. Alfred A. Knopf. New York.
- UN Women (2015). Facts and figures: Leadership and political participation. <http://www.unwomen.org/en/what-we-do/leadership-and-political-participation/facts-and-figures>. Accessed 30 January 2015.
- Wieland, A. and Sarin, R. (2012). Domain specificity of sex differences in competition. *Journal of Economic Behavior & Organization*, 83(1):151–157.
- Wozniak, D., Harbaugh, W. T., and Mayr, U. (2014). The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices. *Journal of Labor Economics*, 32(1):161 – 198.
- Zhang, J. (2013). Can experimental economics explain competitive behavior outside the lab? Unpublished manuscript.

## A Appendix: Additional tables and figures



Figure A1: Screenshot from the online game



Figure A2: Table selection in the online card game community

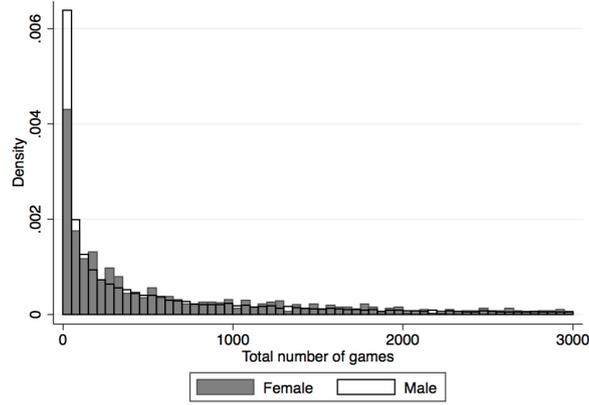


Figure A3: Distribution of total number of games played, by gender (truncated at 3,000)

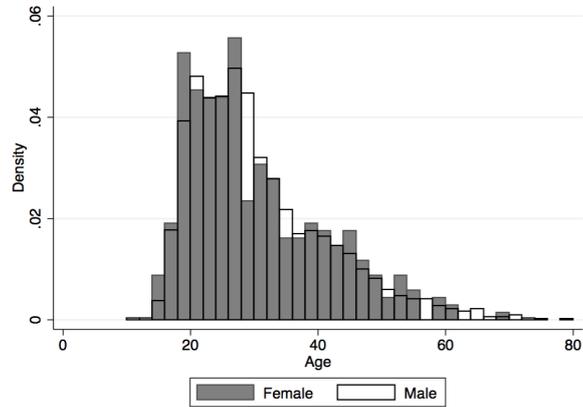


Figure A4: Distribution of age, by gender

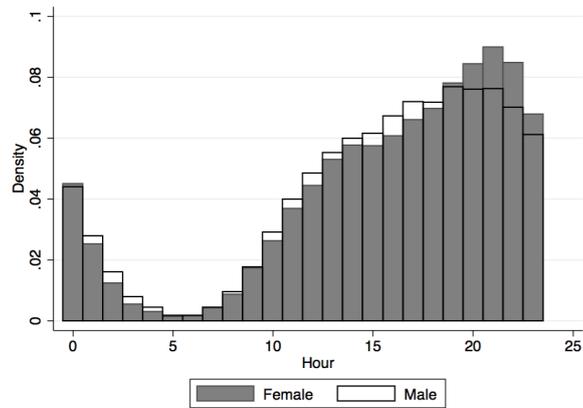


Figure A5: Distribution of the hour of the day when games are played, by gender

Table A1: LIKELIHOOD OF INITIATING A GAME AND RAISING STAKES - Estimated coefficients

VARIABLES	Initiate a game			Increase stakes
	(1)	(2)	(3)	(4)
Female player	-0.134*** (0.012)	-0.139*** (0.013)	-0.117*** (0.012)	-0.218*** (0.028)
Position at table = 2	-0.080*** (0.001)	-0.080*** (0.001)	-0.080*** (0.001)	-0.029*** (0.001)
Position at table = 3	-0.116*** (0.001)	-0.117*** (0.001)	-0.117*** (0.001)	-0.031*** (0.002)
Position at table = 4	-0.133*** (0.002)	-0.135*** (0.002)	-0.135*** (0.002)	-0.023*** (0.001)
Female opponent		-0.013*** (0.002)	-0.014*** (0.002)	-0.018*** (0.003)
Num. opponents knocked = 1		-0.392*** (0.001)	-0.390*** (0.001)	
Num. opponents knocked = 2		-0.698*** (0.002)	-0.694*** (0.002)	
Num. opponents knocked = 3		-0.931*** (0.007)	-0.926*** (0.007)	
Num. games played (log)			-0.019*** (0.002)	0.008 (0.004)
Rank in score = 2			-0.048*** (0.005)	-0.084*** (0.009)
Rank in score = 3			-0.077*** (0.006)	-0.155*** (0.012)
Rank in score = 4			-0.094*** (0.007)	-0.246*** (0.015)
Constant	-0.820*** (0.004)	-0.538*** (0.004)	-0.348*** (0.014)	-0.495*** (0.033)
Observations	16,655,344	16,655,344	16,655,344	16,655,344
Number of players	15232	15232	15232	15232
Pseudo $R^2$	0.00230	0.0312	0.0317	0.00748

Notes: The table displays estimated coefficients from probit models. Dependent variable in columns (1) - (3): initiate any game type; column (4): increase the stakes of the game by “knocking”. Omitted categories for covariates: position at table = 1; female opponent = 0; rank in score = 1; num. opponents knocked = 0. Standard errors are clustered on player ID and are reported in parentheses. \*\*\* p<0.001, \*\* p<0.01, \* p<0.05

Table A2: GENDER INTERACTIONS

VARIABLES	Female opponent (1)	Opponents knocked (2)	Rank (3)	Full model (4)
Female player	-0.115*** (0.011)	-0.115*** (0.012)	-0.083*** (0.020)	-0.079*** (0.020)
Position at table = 2	-0.080*** (0.001)	-0.080*** (0.001)	-0.080*** (0.001)	-0.080*** (0.001)
Position at table = 3	-0.117*** (0.001)	-0.117*** (0.001)	-0.117*** (0.001)	-0.117*** (0.001)
Position at table = 4	-0.135*** (0.002)	-0.135*** (0.002)	-0.135*** (0.002)	-0.135*** (0.002)
Female opponent	-0.014*** (0.002)	-0.014*** (0.002)	-0.014*** (0.002)	-0.013*** (0.002)
Num. opponents knocked = 1	-0.390*** (0.001)	-0.390*** (0.001)	-0.390*** (0.001)	-0.390*** (0.001)
Num. opponents knocked = 2	-0.694*** (0.002)	-0.692*** (0.002)	-0.694*** (0.002)	-0.692*** (0.002)
Num. opponents knocked = 3	-0.926*** (0.007)	-0.922*** (0.008)	-0.926*** (0.007)	-0.922*** (0.008)
Num. games played (log)	-0.019*** (0.002)	-0.019*** (0.002)	-0.018*** (0.002)	-0.018*** (0.002)
Rank in score = 2	-0.048*** (0.005)	-0.048*** (0.005)	-0.048*** (0.005)	-0.048*** (0.005)
Rank in score = 3	-0.077*** (0.006)	-0.077*** (0.006)	-0.073*** (0.006)	-0.073*** (0.006)
Rank in score = 4	-0.094*** (0.007)	-0.094*** (0.007)	-0.084*** (0.007)	-0.084*** (0.007)
Female player * female opponent	-0.006 (0.006)			-0.009 (0.006)
Female player * 1 opp. knocked		-0.001 (0.003)		0.000 (0.003)
Female player * 2 opp. knocked		-0.013* (0.006)		-0.011 (0.006)
Female player * 3 opp. knocked		-0.038 (0.023)		-0.035 (0.023)
Female player * 2nd rank			-0.009 (0.015)	-0.009 (0.015)
Female player * 3rd rank			-0.036 (0.020)	-0.036 (0.020)
Female player * 4th rank			-0.069** (0.023)	-0.069** (0.023)
Constant	-0.348*** (0.014)	-0.348*** (0.014)	-0.355*** (0.014)	-0.356*** (0.014)
Observations	16,655,344	16,655,344	16,655,344	16,655,344
Number of players	15232	15232	15232	15232
Pseudo $R^2$	0.0317	0.0317	0.0318	0.0318

Notes: The table displays estimated coefficients from probit models. Dependent variable: initiate any game type. Omitted categories for covariates: position at table = 1; rank in score = 1; num. opponents knocked = 0. Standard errors are clustered on player ID and are reported in parentheses. \*\*\* p<0.001, \*\* p<0.01, \* p<0.05