IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10587

Choking under Pressure and Gender: Evidence from Professional Tennis

Danny Cohen-Zada
Alex Krumer
Mosi Rosenboim
Offer Moshe Shapir

FEBRUARY 2017

IZA – Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10587

Choking under Pressure and Gender: Evidence from Professional Tennis

Danny Cohen-Zada
Ben-Gurion University of the Negev and IZA

Alex Krumer
SEW, University of St. Gallen

Mosi Rosenboim
Ben-Gurion University of the Negev

Offer Moshe Shapir
New York University Shanghai and Sapir Academic College

FEBRUARY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Choking under Pressure and Gender: Evidence from Professional Tennis

We exploit a unique setting in which two professionals compete in a real-life tennis contest with high monetary rewards in order to assess how men and women respond to competitive pressure. Comparing their performance in low-stakes versus high-stakes situations, we find that men consistently choke under competitive pressure, but with regard to women the results are mixed. Furthermore, even if women show a drop in performance in the more crucial stages of the match, it is in any event about 50% smaller than that of men. These findings are robust to different specifications and estimation strategies.

JEL Classification: J16, J24

Keywords: gender, performance, competitive pressure, tennis, choking

Corresponding author:
Alex Krumer
Swiss Institute for Empirical Economic Research (SEW)
University of St. Gallen
Varnbüelstrasse 14
CH-9000 St. Gallen
Switzerland
E-mail: alexander.krumer@unisg.ch

I. Introduction

Interest in gender differences in the labor market has a long history that continues to this day. Despite the significant increase in female labor force participation in past decades, there is still considerable under-representation of women in high profile jobs as well as wage significant differences between genders.[1] One possible explanation for this gap is that women self-select themselves into lower paid occupations.[2] However, this reasoning is insufficient, as even in the same workplace and with the same level of experience and education, inequality between men and women persists. One other explanation that has been put forth is that women are discriminated against in the workplace (Black and Strahan, 2001; Goldin and Rouse, 2000). For example, Goldin and Rouse (2000) showed that when orchestras revised their audition policies and began to blind audition with a screen that concealed the identity of the candidate from the jury, women were more likely to be advanced and/or hired. A third reasonable possibility that is directly related to the present study is that men respond to pressure better than women.

In general, the link between pressure and performance has received much academic attention over the years. For example, Baumeister (1984) described a negative relationship between performance and incentives – a phenomenon known as "choking under pressure". Similarly, Yerkes and Dodson's results (1908) imply that increased motivation beyond an optimal

[1] In 2015, only 23 of the Fortune 500 companies were run by women (4.6%). See http://fortune.com/2015/06/29/female-ceos-fortune-500-barra/?iid=sr-link9 (last accessed 16.02.2016)

[2] For example, Dohmen and Falk (2011) showed that women are less likely to work in jobs with a variable payment scheme, where salaries are generally higher. Similarly, Kleinjans (2009) found that women have a greater distaste for competition and therefore self-select themselves into less competitive occupation fields, which are generally lower-paid. In addition, Buser, Niederle and Oosterbeek (2014) revealed gender differences in competitiveness that are responsible for gender differences in career choice. These findings are in line with Niederle and Vesterlund (2007) who showed that although women do not perform worse than men on an arithmetic task, they choose to shy away from a competitive compensation scheme whereas men embrace it.

level may harm performance. Also, Dandy, Brewer and Tottman (2001) showed that Australian basketball players' free-throw performance was worse during games than during training. In addition, Dohmen (2008) found that male professional soccer players' performance is negatively affected by the presence of a supportive audience. More recently, Hickman and Metz (2015) showed that higher stakes increase the likelihood to miss a shot on the final hole in professional golf.

However, there is much less evidence on whether there are gender differences in responding to competitive pressure. Moreover, the existing evidence is quite mixed both from experimental and non-experimental studies. For example, using a laboratory experiment, Gneezy, Niederle and Rustichini (2003) observed that as they increased the competitiveness of the environment, men's performance increased, but not women's. This gender difference was documented only when women competed against men but not when they competed in single-sex environments. On the other hand, in another experimental study, Ariely et al. (2009) showed that high monetary rewards can decrease performance in general, but no gender differences were observed in responding to such rewards. Among the non-experimental studies, Jurajda and Münich (2011) examined entrance exams from several universities taken by the same individuals. Two interesting results were observed: First, women do not generally refrain from applying to prestigious, more competitive institutions, and second, men outperform women in entrance exams for such institutions (although no such difference was documented for exams of less competitive schools). Similarly, Shurchkov (2012) observed that while women perform better in a low-pressure verbal exam, they underperformed men in a high-pressure math-based exam. In addition, Ors, Palomino, and Peyrache (2013) observed that men performed better than women in a highly competitive entry exam to a selective French business school (HEC). However, for the same cohorts of candidates, females performed significantly better both in an earlier, less competitive pass/fail type of exam and, once admitted, during the less competitive first year of their studies. On the other hand, Lavy (2013) found that the performance of teachers in a competitive environment was no different between genders, nor did women's performance vary with the gender mix of the teaching staff. Moreover, the performance of women teachers even improved in a competitive environment relative to a non-competitive one. More recently, Jetter and Walker (2015) found that male and female tennis players perform similarly in high-stakes situations.

The purpose of this paper is to shed additional light on how men and women respond to competitive pressure. For this purpose, we use game-level tennis data on all the first sets of all four 2010 Grand Slam tournaments and examine, within each tennis match, whether and to what degree each gender deteriorates or improves at crucial stages of the match.[3] Our analysis is based on 4,127 women's and 4,153 men's tennis games. Since in tennis there is a significant advantage to the server, we define the probability of losing a game on serve as a proxy for the performance of the server and compare men's and women's performance under different levels of competitive pressure.[4] Applying data from highly qualified competitive tennis players in a real tournament setting with strong incentives to win has several advantages. First, it provides a unique opportunity to observe and measure performance as a function of variables such as heterogeneity in abilities and prizes. Second, an unambiguous definition of the importance of each game can be derived according to the tennis scoring system, as based on an estimate of the impact of the game on the probability of winning the match (Klaassen and Magnus, 2001; González-Díaz, Gossner and Rogers, 2012; Paserman, 2010). The fact that there is a clear winner of each point, game, set and

[3] In the data section, we justify why we use only the first set of each match.
[4] On average, a player wins 72.6% of games on serve.

match provides us with high quality information about the importance of stakes and the performance of the players at every juncture of the match. Third, it allows an analysis of how men and women respond to real-world competitive environments.

Comparing within each match the performance of men versus women in low-stakes games versus high-stakes games we find that men consistently choke under competitive pressure; however with regard to women the results are mixed. Furthermore, even if women show a drop in performance in the more crucial stages of the match, it is in any event about 50% smaller than that of men. This finding is robust to several identification strategies and to different measures of competitive pressure.

It is noteworthy that the biological literature on cortisol, often referred to as the "stress hormone" supports our finding. This literature indicates that in response to achievement challenges, cortisol levels increase more rapidly among men than among women, and that high levels can debilitate the mind's critical abilities (Goleman and Boyatzis, 2008). To illustrate, Kirschbaum, Wüst and Hellhammer (1992) described a significantly higher release of cortisol among men compared to women during public speaking and performing mental arithmetic in front of an audience (for similar evidence see also Kudielka et al., 1998). In addition, Stroud, Salovey and Epel (2002) obtained a similar finding during mathematical and verbal challenges. In sport-related environments, Doan et al. (2007) observed that poorer golf performance was associated with higher cortisol levels, and Lautenbach et al. (2014) found a negative significant relationship between salivary cortisol and second serve performance in tennis.

Our paper is most closely related to Paserman (2010) who was the first to use data on Grand Slam tournaments in order to analyze whether men and women respond differently to competitive pressure in a setting with large monetary rewards. He concentrated on how players adjust their behavior during high stakes points and showed that they play more conservatively and less aggressively, thereby making fewer unforced errors and hitting fewer winning shots. He did not directly assess the effect of competitive pressure on the likelihood of winning the match but rather relied on a game-theoretical model to calculate this probability. Interestingly, similar to our finding, his calculations led him to conclude that men's performance deteriorates at least as much as women's on more important points.

We build on his paper but expand on it by directly analyzing the effect of competitive pressure on the probability of losing a game on serve. Our approach is quite different because instead of concentrating on how players adjust their strategic behavior during high stakes games relative to low stakes ones, we directly examine whether players ultimately win more often or less frequently in high stakes games. This issue is especially important, because the adjusted behavior of the players may affect the likelihood of their winning the game, regardless of the number of winning shots and unforced errors they make. For example, playing more conservatively and less aggressively may also increase the number of forced errors. In addition, it may affect not only the player's number of winning shots and unforced errors but also those of his opponent. Thus, it is not at all clear how changes in strategic behavior ultimately affect the probability to win the game. Moreover, we argue that if a player decides to play less aggressively at crucial stages of the match and this choice led him to winning games at higher rates relative to non-pressured games, it means that there was an improvement under pressure. In contrast, if this same choice led to losing games at higher rates then it means that there was choking under pressure. For this reason, in order to determine which gender chokes more under pressure it could be more advantages to examine the effect of that competitive pressure on the likelihood of winning the game rather than on the number of unforced errors or winning shots. Another advantage of our measure is that it is completely objective. While two observers can debate whether or not a certain shot should be considered as a forced or unforced error, and whether or not the previous shot led to the forced error, in our case winning a game is an undeniable fact.

Although our finding is interesting, caution must be exercised in generalizing it to other professions where decisions are made under high-pressure environments. For one thing, in our setting players compete in a single-sex contest and it is possible that women would respond differently to competitive pressure in mixed-sex contests. In fact, Gneezy, Niederle and Rustichini (2003) found that while men perform similarly in both contests, women's performance is significantly higher in a single-sex contest. Thus, in places where women are required to compete with men (such as in the labor market), they might respond worse to competitive pressure. Second, many other differences exist between professional athletes and other professionals (lawyers, doctors, economists, brokers) such as intensified use of motors skills, frequent changes of location, and being far from home and family. Therefore, while women may respond to pressure better than men in tennis, this may not necessarily carry over to other areas. Nevertheless, our robust evidence that women can respond better than men to competitive pressure is compelling and should spark further investigation of this issue.

The rest of the paper is organized as follows: Section 2 explains the basic rules of tennis. Our theoretical model is presented in Section 3. The data are described in Section 4. In Section 5 we present the estimation strategy. Section 6 reports the results and Section 7 offers concluding remarks.

2. Basic Rules of Tennis

A tennis match is played between two players who stand on opposite sides of a net. One player is designated as the server, and the other as the receiver. Service alternates game by game between the two players. The first player to win four single points wins the game, provided that he or she also has at least two points more than his or her opponent. Typically, a player wins a set by winning at least six games and at least two games more than the opponent. If one player has won six games and the opponent five, an additional game is played. If the player who was in the lead wins that game, he or she wins the set 7–5. On the other hand, if the player who was trailing wins the game, a tiebreak is played.[5] In that case, the set is decided by the player who wins at least seven points in the tiebreak, provided he or she also has at least two points more than his or her opponent. A match is won when a player wins the majority of sets. In Grand Slam tournaments women always play in the form of best-of-three sets, whereas men play in the form of best-of-five sets.

3. The Model

Consider a tennis game played between a server and a receiver, both of the same gender. We denote the probability that the server loses game g of match m to the receiver by, $\pi Break_{gm}$, and assume this probability to be a negative linear function of the performance of the server, PS_{gm}, and a positive linear function of the performance of the receiver, PR_{gm}. Formally, the probability that the server loses game g of match m to the receiver is given by

$$\pi Break_{gm} = \beta_0 + \beta_R \cdot PR_{gm} - \beta_S \cdot PS_{gm} \quad ; \quad \beta_S > \beta_R \geq 0 \quad , \tag{1}$$

[5] A tiebreak is never played in the deciding set in the three Grand Slam tournaments - the Australian Open, the French Open, and Wimbledon.

where β_R and β_S refer to the effects of the receiver's and the server's performance, respectively, on the probability to lose a game on serve. As indicated in equation (1), we also assume that this probability depends more strongly on the performance of the server rather than the receiver ($\beta_S > \beta_R$). This assumption is easily justified as the serve is the only shot in tennis where the server has complete control over where he places the ball, how to spin it, and how fast he hits it. Therefore, if the server serves sufficiently well, the receiver will have almost no chance to win the point. This situation roughly resembles that of penalty shots in soccer where since the kicker has control of where and how to kick the ball, the common perception is that success depends more on him than on the goalkeeper. Indeed, our data imply that on average a player wins 72.6% of games on serve. This number is very similar to the success rate in penalty shots (74.2%, Dohmen 2008), which implies that the advantage of the server over the receiver is very similar to the advantage of the kicker over the goalkeeper. In fact, Apesteguia and Palacios-Huerta (2010) conducted a survey among 240 Spanish players and coaches indicating that not even a single player considered the performance of the goalkeeper to be more important than the performance of the kicker in penalty kicks.

We posit that the performance of each player consists of two components. One is that each player has a basic performance level that is a linear function of his attributes (such as his ranking, body mass index (BMI), height, and whether he has a home advantage or not).[6] We denote the vectors of the server's and receiver's attributes by S_{gm} and R_{gm}, respectively. These vectors have an index *gm* because the identity of the server and receiver alternates game by game within a given match. The other component is that the performance of both the server and the receiver are

[6] See Krumer, Rosenboim and Shapir (2016) and Koning (2011) who, respectively, described the effect of physical characteristics and home advantage in professional tennis.

influenced by the level of competitive pressure they face. We allow men ($W_j = 0$) and women ($W_j = 1$) to respond differently to pressure, but impose that for a given gender there is no systematic difference in how servers and receivers respond to competitive pressure.

Formally, the performance of the server and receiver are given by

$$PS_{gm} = S_{gm}^{'} \cdot \delta_S - \gamma_M \cdot Pressure_{gm} - \gamma_D \cdot W_m \cdot Pressure_{gm} \qquad (2a)$$

and

$$PR_{gm} = R_{gm}^{'} \cdot \delta_R - \gamma_M \cdot Pressure_{gm} - \gamma_D \cdot W_m \cdot Pressure_{gm} \qquad (2b)$$

In equations (2a) and (2b), the parameters δ_S and δ_R measure the effect of the attributes of the server and the receiver on their corresponding level of performance, the parameter γ_M measures the effect of competitive pressure on men's performance, and the expression $\gamma_M + \gamma_D$ measures this effect on women's performance. A negative value of γ_M indicates that men's players performance improves under pressure, while a positive γ_M implies that pressure causes them to choke. Similarly, the parameter γ_D reflects the differential effect of pressure among men and women, where a positive value of γ_D implies that women respond to pressure worse than men, while a negative γ_D implies that they respond better.

Substituting the performance of the server and the receiver from equations (2a) and (2b) into equation (1) and rearranging terms yields

$$\pi Break_{gm} = \beta_0 + R_{gm}^{'} \cdot \delta_R \cdot \beta_R - S_{gm}^{'} \cdot \delta_S \cdot \beta_S + (\beta_S - \beta_R) \cdot \gamma_M \cdot Pressure_{gm} + (\beta_S - \beta_R) \cdot \gamma_D \cdot W_m \cdot Pressure_{gm} \qquad (3)$$

Obviously, the performance of both players is also a function of a set of observed and unobserved factors that do not vary within each match such as surface, weather, specific matchup between the two players, etc. In addition, the average probability to lose a game on serve is different among men and women regardless of the level of competitive pressure. This is true especially because, due to physiology differences, men on average serve much more powerfully than women. As such, physiological differences between men and women do not vary within each match, so that we can control for them by including match fixed effects, μ_m in equation (3). By doing so, like in any "difference in difference" approach, we acknowledge that the probability to lose a game on serve is different between men and women and thus identify the effect of competitive pressure by comparing the *difference* in the probability to lose a game on serve between low stakes and high stakes games among men and women. This difference-in-difference estimator would not be biased with regard to any difference between men and women that is fixed within each match.

Since the identity of the server and receiver alternate game by game within each match, in addition to match fixed effects, we are able to include in the estimation also control variables for either the characteristics of the server or the receiver but not both.[7] Thus, we omit the characteristics of the receiver and estimate the following equation:

$$\pi Break_{gm} = \beta_0 - S_{gm}^{'} \cdot \delta_S \cdot \beta_S + (\beta_S - \beta_R) \cdot \gamma_M \cdot Pressure_{gm} + (\beta_S - \beta_R) \cdot \gamma_D \cdot W_m \cdot Pressure_{gm} + \mu_m \qquad (4)$$

[7] To understand this, suppose that the height of the player who serves the first game is 1.80 meters, while the height of the second player is 1.90 meters. In this case, in the second game the height of the server would be 1.90, while the height of the receiver is 1.80. Similarly, these two numbers continue to switch in every game. Thus, since the two variables sum up to a constant number, to avoid multicollinearity we must omit one of them.

According to this equation, the pressure coefficient, $(\beta_S - \beta_R) \cdot \gamma_M$, measures how pressure affects the probability to lose a game on serve among men ($W_j = 0$), while the interaction coefficient between $Pressure_{gm}$ and W_m, $(\beta_S - \beta_R) \cdot \gamma_D$, measures the differential effect of pressure between men and women. It is then straightforward that under our reasonable assumption that $\beta_S - \beta_R$ is positive, the sign of the coefficient of pressure is the same as the sign of γ_M. Put formally,

$$sign\,[(\beta_S - \beta_R) \cdot \gamma_M] = sign\,[\gamma_M] \tag{5}$$

Similarly, the sign of the interaction coefficient between *Pressure* and *Women* is the same as the sign of γ_D. Put formally,

$$sign\,[(\beta_S - \beta_R) \cdot \gamma_W] = sign\,[\gamma_D] \tag{6}$$

Thus, by estimating equation (3) we can properly identify the sign of γ_M and γ_D. It is noteworthy that our identification of the signs of γ_M and γ_D do not rest on the assumption that the probability to lose a game on serve is a measure of the performance of the server only (i.e., that $\beta_R = 0$). The signs of γ_M and γ_D are properly identified even if $\beta_R > 0$ as long as $\beta_S - \beta_R > 0$. In this estimation, according to equation (5) a negative coefficient of $Pressure_{gm}$, $(\beta_S - \beta_R) \cdot \gamma_M$, implies that men improve under pressure, while a positive coefficient indicates that men choke under pressure. Note that since the coefficient of the pressure variable is $(\beta_S - \beta_R) \cdot \gamma_M$, and β_S and β_R are unobserved, we can identify only the sign of γ_M but not its size.

Proposition 1: A positive pressure coefficient on the probability to lose a game on serve implies that men choke under pressure while a negative coefficient implies that men improve under pressure.

It is noteworthy that a positive coefficient of $Pressure_{gm}$ can be interpreted to result from clutching ($\gamma_M < 0$) only if one is ready to assume that the probability to lose a game on serve depends more strongly on the performance of the receiver rather than the server ($\beta_S - \beta_R < 0$). As we already showed earlier that this is an unreasonable assumption, in what follows we ignore this interpretation.

Similarly, from equation (6) it is straightforward that:

Proposition 2: A positive interaction coefficient between W_m and $Pressure_{gm}$ implies that women respond to pressure worse than men, while a negative coefficient implies that they respond better than men.

In summary, by estimating equation (3) we can now evaluate whether men and women choke under pressure and whether their response to pressure is significantly different. This is done in Sections 5 and 6.

4. Data

We collected data on all the matches of the four 2010 Grand Slam singles tournaments (the Australian Open, the French Open, Wimbledon and the US Open). We concentrate specifically on

Grand Slam tournaments, because they are the only tournaments in which the prizes are equal for men and women. This choice is particularly essential given the large body of literature showing that prizes and incentives are strongly related to performance (Lazear, 2000; Dohmen and Falk, 2011, among many others). An alternative option could have been to concentrate on top tier ATP and WTA tournaments, in which the advantage is that both men and women play in the form of best-of-three sets. However, a major disadvantage is that in ATP and WTA tournaments not only are the prizes not equal for the genders, but they are systematically larger for men than for women. Thus, this factor could have led us to conclude that women respond better to pressure even if they actually do not. That is, if we used these tournaments, one could argue that our finding that women respond better to pressure is driven by the fact that they simply face less pressure because they are competing for smaller prizes. In stark contrast, the fact that women play best-of-three sets while men play best-of-five sets only works against our finding that women respond better to competitive pressure. For one thing, playing only three sets makes the first set more crucial for winning the entire match, thus putting more pressure on women relative to men. In other words, for a given score in the first set, women face greater pressure than men, because, unlike men, if they lose the first set they must win the next set in order to stay in the match. This argument is confirmed by our data as well. In Grand Slam tournaments, the probability of winning the match, conditional on losing the first set, is approximately 22.4% and 14.6% for men and women, respectively. Moreover, as we show below, our pressure index does take into account the fact the men compete in the form of best-of-five sets, while women play in the best-of-three sets. In fact, possibly because of these reasons, previous studies that dealt with gender differences in tennis matches also used data on Grand Slam tournaments (Paserman, 2010; Jetter and Walker, 2015).

For each match we have information available on the names of the players, the round of the tournament, the prize for winning the match, the surface of the courts, the players' height and BMI, dummy variables for whether each of the players has a home advantage, and the players' 52-week ranking prior to the beginning of the tournament. This ranking takes into account all of the results in all of the professional tournaments that took place over the past 52 weeks. Following Klaassen and Magnus (2001), we rescaled the ranking of each player as $8-\log_2(Rank)$ and denoted the rescaled indexes of the first server and second server as *FSRI* (first server ranking index) and *SSRI* (second server ranking index), respectively. These indexes serve as measures of the players' abilities, with a higher index indicating a player's better ability. We use these ranking indexes in order to construct our pressure measures and also a measure for the relative strength of the first server, defined as *Gaprank* = *FSRI* – *SSRI*.

In addition, we have game-by-game data including information on who served in each game and who won the game. The data were collected from official tournament web sites, betting and news sites among others (see Table A1 in the Appendix for the full list). Our database covers 508 men's and 508 women's matches. Unfortunately, 26 men's and 7 women's matches ended due to injury, 4 matches of each gender lacked information on the BMI of the players, and another 48 men's and 41 women's matches lacked information on the distribution of breaks during a set, leaving us with a total of 430 men's and 456 women's matches which contain 4,153 men's and 4,127 women's games.

A potential concern is that these incomplete observations are not missing completely at random, but rather predominantly refer to lower quality players, which may introduce endogenous selection into the sample. To obviate this concern and show that the missing values are randomly distributed across all observations, we used the following two-step procedure. First, we partitioned the data into two parts, where one set contains the matches with the missing values (the omitted matches) and the other the matches with the non-missing values (the included matches). In Table A2 in the Appendix we report the average value of each player's characteristics and in parentheses their standard deviation separately for each gender and each dataset. Columns 1 and 4 refer to the omitted matches for men and women, respectively, while Columns 2 and 5 refer to the included matches for men and women, respectively. Then, separately for each gender, we run a set of univariate regressions of each of the variables presented in Table A2 on a dummy variable indicating whether the specific observation was omitted or included in the sample. The coefficient of this dummy variable and its standard error are presented in Columns 3 and 6 of Table A2, for men and women, respectively. The results show that none of the players' characteristics differ significantly between the two datasets, which indicates that the incomplete observations are missing at random, and thus selection into the sample is not a concern.

We use data on only the first set of each match for several reasons. First, it avoids the possible influence of asymmetry and fatigue in the following sets. Second, winning the first set provides a huge advantage for winning the entire match. In fact, our data suggests that among women 85.4% of those who won the first set also won the entire match. The corresponding number among men is 77.6%. Third, the performance in the following sets may be affected by different psychological or strategic momentum effects (Malueg and Yates, 2010; De Paula and Scoppa, 2015).

Our dependent variable, $\pi Break_{gm}$, is a dummy variable that gets the value of 1 if the server lost the game on serve and zero otherwise. Table 1 presents the descriptive statistics. We can see that in about 27.4% of the games the server lost the game on serve. The corresponding number among men is 20% while among women it is 34.9%.

As our analysis focuses on the differential effect of competitive pressure on performance, one of our key independent variables is a dummy, $Wome_i$, that gets the value of 1 if the game is a women's game and zero otherwise.

We use three different measures for competitive pressure. Two are based on the importance of the game for winning the entire match. Following Morris (1977), Klaassen and Magnus (2001) and Passerman (2010), we define the importance of the game as the difference in the probability of winning the entire match as a result of winning or losing the current game. The first measure, $PressureGR_{gm}$, takes into account that the importance of each game differs by gender and by the rankings of the players. Thus, it is constructed separately for each gender using the following four-step procedure: First, for each score, we run a probit regression of the probability of the first server to win the first set as a function of the rescaled ranking indexes of the two players. For example, estimating the probability of the first server to win the first set conditional on leading 2:0 among women yields:

$$\pi WinSet|_{2:0,\ women=1} = \Phi(1.282 + 0.149FSRI - 0.165SSRI) \qquad (7)$$

where *FSRI* and *SSRI* denote the above-defined rescaled ranking indexes of the first server and the second server, respectively. Similarly, estimating the probability of the first server to win the first set from a score of 1:1 yields:

$$\pi WinSet|_{1:1,women=1} = \Phi(-0.077 + 0.234FSRI - 0.147SSRI) \qquad (8)$$

Second, for each game g of match m we calculated the importance of the game for winning the entire set as the difference in the probability of winning the first set as a result of winning or losing the current game. For example, the importance of a game played at a score 1:0 (in favor of the first server) for winning the first set is

$$ImportanceGR_{gm}|_{1:0,women=1} = \pi WinSet|_{2:0,women=1} - \pi WinSet|_{1:1,women=1} = \Phi(1.282 + 0.149FSRI - 0.165SSRI) - \Phi(-0.077 + 0.234FSRI - 0.147SSRI) \quad (9)$$

Third, we took all the matches in which the first server won the first set and only among these matches estimated a probit regression of the probability to win the entire match as a function of the rankings of the two players. For women, this estimation yielded:

$$\pi WinMatch|_{winset=1,women=1} = \Phi(0.956 + 0.151FSRI - 0.129SSRI) \quad (10)$$

Obviously, this conditional probability is, on average, much larger for women than for men because women play in the form of best-of-three sets while men in best-of-five sets. Finally, for each game g of match m we calculated the level of competitive pressure as the multiple between the importance of the game for winning the first set (calculated in step 2) and the probability to win the entire match conditional on winning the first set (calculated in step 3). For example, the value of $PressureGR_{gm}|_{1:0,women=1}$ is obtained by multiplying the expression in equation (9) by that in (10). The average values of this measure by gender and set status are reported in Table 2. We can see that for most scores the importance of the game is larger for women than for men, which stems from the fact that they play under best of three sets while men under best of five sets.

One nice property of this measure of competitive pressure is that the importance of a game from the perspective of the server is exactly identical to that of the receiver.[8] Another noteworthy feature of this measure is that it evolves very non-linearly within a given match, generating sufficient variation that can be used for the estimation.

Our second measure of competitive pressure is taken from Paserman (2010) who constructed it using data on all Grand Slam tournaments during the years 2006-2007. Although the measure is based on the importance of each point rather than the game, in Table 3 he reports the average values of this measure by set status. We use these average values. Thus, this measure does not vary by gender or by ranking. In our study it is denoted by *PressureBP* and its values are reported in Table 2. One advantage of using this measure is that since it was constructed from data on other tournaments, it is not a function of the outcomes in our sample. Also, it is presumably more exogenous than $PressureGR$ as it was not derived as a function of rankings and gender.

A third measure for competitive pressure, *PressureD*, is one that is dichotomous. Accordingly, a game is defined as a high-stakes game only if each of the players previously won at least four games (i.e., only if the score is 4:4, 4:5, 5:4, 5:5, 6:5, 5:6). Otherwise, the game is considered as a low-stakes game and thus gets a value of zero. In all these instances, either both players can win the set by winning not more than two subsequent games or one of them can win the set by winning only one game and the other by winning three subsequent games. This measure is actually a dichotomization of *PressureBP* as, according to this measure, the pressure level in each of these scores is higher than the pressure level of any other score (see Figure 1).

[8] There is however a disadvantage, namely, that it does not take into account that the value of winning may be different for the two players. For example, this would be the case if a win would have a differential effect on players' rankings. Similar to Malueg and Yates (2010) we only acknowledge this possibility but see no practical way to control for it.

5. Estimation Strategy

Studying the effect of competitive pressure on performance among men and women is quite challenging. A naïve approach of correlating measures of competitive pressure and performance will yield biased and inconsistent estimates because unobserved determinants of performance are likely to be correlated with competitive pressure. For example, the level of competitive pressure is on average higher when the score is tighter, which generally occurs when the asymmetry between the players' abilities is smaller. Thus, any unobserved characteristic of the players that is associated with the tightness of the match can render such a naïve approach invalid. Hence, we use several different strategies in order to tease out the effect of selection. First, we include in all our estimations a set of match fixed effects in order to control for any unobserved factor that is fixed within a match such as weather, temperature, surface, the specific matchup between the two players, and any difference between men and women that is fixed within a given match. Thus, like in our theoretical model, our basic specification takes the following form

$$\pi Break_{gm} = \alpha_1 \cdot PressureGR_{gm} + \alpha_2 \cdot W_m \cdot PressureGR_{gm} + \beta_1 \cdot X_{gm} + \mu_m + \varepsilon_{gm} \qquad (11)$$

where the dependent variable, $\pi Break_{gm}$, is the probability of the server to lose game *g* of match *m* on serve; $PressureGR_{gm}$ is our central measure for the level of competitive pressure in game *g* of match *m*; W_m is a dummy variable indicating whether match *m* is a women's match (*Women* = 1); X_{gm} is a set of characteristics of the server in game *g* of match *m* (such as his relative ranking measured as the difference between his and the receiver's ranking, a dummy for whether he has a home advantage, and his height and BMI). This "difference-in-difference" specification allows us to compare the performance of men versus women in low-pressure games versus high-pressure one. According to the propositions of our theoretical model, a positive value of α_1 implies that the server chokes under pressure, while a negative value implies that he clutches under pressure. Similarly, a positive coefficient of α_2 implies that women respond to pressure worse than men while a negative value implies that they respond better.

Next, as the identity of the server and receiver alternates game by game within a given match (where the first server serves in all the games with an odd serial number, while the first receiver serves in all the games with an even serial number), the characteristics of the server and receiver also alternate every game within a given match. In this case, including match fixed effects in the estimation is not sufficient to control for any unobserved characteristics of the server and receiver that might affect performance. Therefore, we next add to our basic specification interaction terms between the match fixed effects and a dummy for whether the serial number of game is odd (OG_{gm}). Thus, for each match we now have two fixed effects, where one is for all the odd games and the other for all the even games, which actually eliminates the need to control for the set of characteristics of the players as they are held constant within each fixed effect. Formally, this specification takes the following form:

$$\pi Break_{gm} = \alpha_1 \cdot \Pr essureGR_{gm} + \alpha_2 \cdot W_m \cdot \Pr essureGR_{gm} + \mu_m + \mu_m \cdot OG_{gm} + \varepsilon_{gm} \qquad (12)$$

Such a specification allows us to compare the performance of men versus women in low-pressure games versus high-pressure games only within the games of a given match in which the identity of the server and the receiver are kept constant and any unobserved characteristic of the players is perfectly controlled for. We therefore view these estimates as our preferred results.

We first estimate all these equations using a fixed effect linear probability model (LPM). We prefer LPM as our main results because of three main reasons.

First, while this model uses all of the observations, a fixed effect logit model can only use observations for which the outcome variable varies within each match. Thus, it omits matches in which the outcome is fixed, which may bias the estimates.[9] This is especially true in our specific context in which the outcome is generally fixed within a match when the server consistently wins each game on serve (i.e., when the server doesn't choke under pressure). Thus, omitting these observations introduces selection into the treatment. Second, any non-linear estimation such as logit relies on the functional form while in the linear model the fixed effects account for variation in the data in a completely general way. Third, a logit fixed effect model yields consistent estimates only under the stronger assumption of strict exogeneity, while LPM requires exogeneity to hold only within a fixed effect.[10] Nevertheless, although we prefer using the LPM fixed effect model

[9] In logit and probit fixed effect models $\Pr\left[\pi Break_{gm}=1\right]=G\left[\alpha_1\cdot \Pr essureGR_{gm}+\alpha_2\cdot W_m\cdot \Pr essureGR_{gm}+\beta_1\cdot X_{gm}+\mu_m\right]$ where $G(.)$ is the cumulative density function (CDF) for either the standard normal or the logistic distribution. It is easy to see that unlike in a linear fixed effect model, because of the function $G(.)$ we cannot eliminate μ_m by using within transformation. Moreover, if we attempt to estimate μ_m directly by adding dummy variables for each match, the estimates of μ_m are inconsistent, and unlike the linear model, the estimates of α_1 and α_2 becomes inconsistent too. Thus, the only way to obtain consistent estimates for α_1 and α_2 is to eliminate μ_m from the equation. In a probit model this is completely impossible. Thus, including match fixed effects in a probit model will yield biased estimates due to the well-known Incidental Parameter Problem (Neyman and Scott, 1948; Greene 2004). Unlike probit models, the logit functional form does enable us to eliminate μ_m from the equation but only under the assumption that the dependent variable changes between the games within a given match. For this reason, it drops matches in which the outcome is fixed within a given match.

[10] Although LPM has the disadvantage that it produces predicted probabilities outside the range 0-1. As Wooldridge (2002) argues, "If the main purpose is to estimate the partial effect of [the independent variable] on the response probability, averaged across the distribution of [the independent variable], then the fact that some predicted values are outside the unit interval may not be very important" (p. 455).

for our main results, we also test the sensitivity of our estimates to using either a logit fixed effect model or GLS.

6. Results

6.1 Main results

Column 1 of Table 3 presents the results from estimating equation (11), where standard errors clustered at the match level are in parentheses. These standard errors correct for both serial correlation and heteroscedasticity. The results show that the coefficient of the pressure variable is positive and significant which implies that men choke under pressure. A one standard deviation increase in our pressure index increases the probability to lose a game on serve by 4.9 percentage points, which is about 25% of the sample mean (the mean value of the probability of break among men is 0.2). In addition, the interaction coefficient α_2 is negative and significant at the 10% level which implies that women choke under pressure less than men. Among women, a one standard deviation increase in our pressure index increases the probability of losing a game on serve by only 2.8 percentage points, which is about 8% of the women's sample mean.

To show that our results are insensitive to the set of controls included in the estimation we also estimated the same equation without controlling for any of the characteristics of the server. The results, reported in Column 2 of Table 3, are very similar to those from our basic specification, and the size of the pressure effect is almost unchanged.

A common threat to the validity of any difference-in-difference analysis of this nature is underlying trends in the data. In other words, the causal interpretation of our finding rests on the identifying assumption that other than pressure there was no other factor that changed across games and affected the two genders differently. We therefore use several specifications to support this assumption. In Columns 3 and 4 of Table 3 we control for any factor that changes smoothly across games (such as fatigue) by adding to our basic specification a linear game trend and a quadratic time trend, respectively. Finally, in Column 5 we include gender-specific game trends in the estimation. All these specifications yield very similar results relative to our basic specification. The coefficient of pressure is always positive and significant and very similar in size. Similarly, the interaction coefficient is always negative and similar in size and significant at the 10% level (except for the last specification, presented in Column 5, where the p-value is slightly above 10%).

Another concern is that the performance of the players is also a function of the score at which the game is played. For example, a player may perform better when he leads a set relative to when he is behind in a set. In addition, if he is far behind he may decide to stop making substantial efforts in order to save energy for the next set. For this reason, it is important to estimate the same specification but only for games in which the status of both players is symmetric in order to minimize strategic considerations. Thus, in Column 6 of Table 3 we estimate the same specification but only for tie games.[11] The results show that while the main pressure effect is only slightly smaller, the interaction term becomes much larger, which implies that women respond substantially better than men to competitive pressure (significant at the 1% level). In fact, this estimation indicates that women do not choke at all under pressure.

Although, as already mentioned, we prefer the LPM estimates as our main results, we also test the sensitivity of our estimates to using either a logit fixed effect model or GLS. The logit estimates are reported in panel B of Table 3 and the GLS in Table 3A in the Appendix. Both the

---

[11] More specifically, we consider the games that are played when the score is 0:0, 1:1, 2:2, 3:3, 4:4 and 5:5.

logit and GLS estimates even strengthen our main finding that women respond better to pressure. In addition, in all these specifications, the interaction term is always negative and significant at the 1% level.[12]

Two related concerns exist regarding the distribution of the rankings of the two genders. First, one may argue that our estimation does not take into account that the ranking distribution is different among the genders, which may affect the results. Second, even if the average rankings and the standard deviation of the rankings are identical for the two genders, it can still be the case that the difference between a top player and a mediocre player is small for men but large for women. We use two strategies to eliminate these concerns. First, we re-estimate our basic specification in Column 2 but instead of using the variable *gaprank* as a control measure for the relative strength of the server, we normalize this variable by gender and use instead the normalized measure. The results, reported in Column 7, indicate that the estimates are almost unaffected by this change. We also tested the sensitivity of the other specifications of Table 3 to replacing *gaprank* with its normalized measure. We consistently find that the estimates are almost unaffected. In addition, in order to test whether *gaprank* has a different effect on the probability to lose a game on serve among men and women we ran a regression of the probability to lose a game on serve as a function of *gaprank* and its interaction with gender, controlling for our entire set of controls, and tested whether the interaction is significant. The results indicate that the interaction

[12] Applying GLS in cases where some of the fitted values are outside the unit interval requires their adjustments. Specifically, we need to set each fitted value that is greater than one equal to 0.999 and each fitted value that is lower than zero equal to 0.001. However, since these adjustments can affect the results, Wooldridge (2006) suggest that it is probably best to abandon WLS and report OLS estimates with heteroscedasticity-robust standard errors (Wooldridge, 2006, p. 292). For this reason, we prefer the OLS as our main results.

term was insignificant, which implies that rankings do not affect outcomes differently among the two genders.[13]

In order to compare the performance of men versus women in low-pressure games versus high-pressure ones only within the games of a given match, in which the identity of the server and the receiver are kept constant and any unobserved characteristic of the players is perfectly controlled for, we next estimate equation (12). The results, reported in Table 4, indicate that in all the specifications the interaction term between pressure and gender becomes larger and much more significant (at the 1% level) relative to the results in Table 3. In addition, the size of the pressure effect is very similar in all the different specifications. Among men, a one standard deviation increase in our pressure index increases the probability of the server to lose a game on serve by 9.0 percentage points, which is a sizeable effect. Among women, it increases the probability to lose a game on serve by only 4.6 percentage points. This result serves as additional evidence that women choke under pressure significantly and substantially less than men. Also, in Panel B of Table 4, we test the sensitivity of these estimates to using a logit fixed effect model. Again, the results indicate that the logit estimates even strengthen our main finding that women respond better to pressure. In addition, in all these specifications, the interaction term is always significant at the 1% level.

Finally, the fact that our pressure index increases very sharply after game number eight (see Figure 2) allows us to use a "regression discontinuity" approach in which we estimate the effect of competitive pressure only within a discontinuity sample that includes three games from each side of the discontinuity. Specifically, we concentrate on games 6-12 and exclude data on

[13] These estimations are not presented due to space limitations but are available upon request.

games 1-5. The results, reported in Columns 5 and 6, indicate that the coefficient of the interaction term between *PressureGR* and *Women* is negative and significant, which again implies that women respond better to competitive pressure relative to men.

In Table 5 we report separate regressions for men and women. Columns 1 and 2 include only match fixed effects in the estimation, while Column 3 and 4 also include interactions between the match fixed effects and a dummy indicating an odd-numbered game. The results indicate that the effect of competitive pressure on the probability of the server to lose a game on serve among women is between 50%-60% of that among men. When we estimate the equations only among tie games the results indicate that while men choke under pressure women do not choke at all (Columns 5 and 6).

6.2 Robustness analysis

One may be concerned that since our pressure index was constructed as a function of gender and rankings it may be endogenous. Although we do not think that this is a valid concern as we do control for these observed factors in our estimation, we next use *PressureBP*, which is not a function of gender and rankings, as an instrumental variable for our *PressureGR* measure. It is presumably more exogenous and also strongly correlates with our pressure measure (the raw correlation between the two measures is 0.62). Thus, it can serve as a valid instrument for estimating the effect of competitive pressure on performance. In addition, this measure also has the advantage of being constructed from data on other tournaments and thus is not a function of the outcome in our sample. The results are reported in Table 6. Again, the estimates indicate that men choke under pressure more than women. In one specification the size of the pressure effect among women is less than one half of the effect among men, and in the other specification women do not choke under pressure at all.

As the last step, we test the robustness of our findings by using a different pressure measure, *PressureD*. According to this dichotomous measure, a game is considered a high-stakes game only if each of the players has already won at least four games.[14] The results, reported in Columns 5 and 6 of Table 6, once again indicate that men do choke under pressure, and interestingly it can be observed that women do not choke at all in crucial stages of the match. Taken together, using several estimation strategies and several measures for competitive pressure we always obtain the same finding that men choke under pressure more than women.

7. Concluding Remarks

In this paper we used a real tournament setting with large monetary rewards in order to examine whether professional men and women choke under competitive pressure and if their response to competitive pressure is different. Based on our analysis of 8,280 men's and women's tennis games we find that men consistently choke under competitive pressure, but with regard to women the results are mixed. Furthermore, we establish that even if women show a drop in performance in the more crucial stages of the match, it is in any event about 50% lower than that of men. One implication of this finding is that, especially for men, incentives might actually reduce performance, because they increase the level of the stakes. Thus, our results do not seem to support the claim that gender differences in wages in the labor market can be attributed to the fact that women respond more poorly to competitive pressure.

[14] More specifically we consider the games that are played when the score is 4:4, 5:4, 4:5, 5:5, 6:5 and 5:6 as high stakes games (the first number is the server's score and the second number is the receiver's score).

Our finding about women's superiority in responding to competitive pressure is consistent with evidence in the biological literature that the levels of cortisol, which is known to impede the performance of both men and women, commonly escalate more substantially among men than women in response to achievement-related challenges.

Our finding is also in line with the literature showing that women also respond similarly or even less strongly to other psychological effects when they compete in single-sex tournaments. For example, Gauriot and Page (2014) found evidence of a momentum effect among men but not among women. They showed that winning a point when the ball bounced very close to the court's line increases the probability to win the next point, but only among men. Similarly, Cohen-Zada, Krumer and Shtudiner (2017) recently observed that in professional judo competitions, men experience psychological momentum, whereas women do not. Jetter and Walker (2015) found mixed results regarding psychological effects on men and women. When they investigated whether men and women respond differently to psychological momentum in tennis by looking at how an additional win in the most recent ten matches increased the probability of winning, they observed no differences between men and women. However, they also found that top women players performed slightly better than men in the most important tournaments relative to the less important ones. Similarly, De Paula and Scoppa (2015) found no gender differences in psychological momentum when examining whether winning the second set in tennis, relative to winning the first one, affects the probability of winning the third set.

Still, caution must be exercised in generalizing the findings of this study to the labor market. For one thing, while we analyzed how female tennis players respond to pressure in a contest that is homogeneous with regard to gender, in the labor market women are required to respond to competitive pressure in a different setting where they compete with men. In fact, several studies find the men respond better than women to competitive pressure when women have to compete with men rather than between themselves (Niederle and Vesterlund, 2010; Niederle and Vesterlund, 2011; Niederle, 2014). In addition, tennis players may have different preferences and characteristics which may not necessarily make them a representative subject. Nonetheless, the fact that we have uncovered such robust evidence that women can respond better than men to competitive pressure calls for further investigation in other real-life tournament settings.

References

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

Apesteguia, J., & Palacios-Huerta, I. (2010). Psychological pressure in competitive environments: Evidence from a randomized natural experiment. *The American Economic Review*, *100*(5), 2548-2564.

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*(3), 610-620.

Black, S. E., & Strahan, P. E. (2001). The division of spoils: rent-sharing and discrimination in a regulated industry. *American Economic Review*, 91(4), 814-831.

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, competitiveness and career choices. *The Quarterly Journal of Economics*, *129*(3), 1409-1447.

Cohen-Zada, D., Krumer, A., & Shtudiner, Z. (2017). Psychological momentum and gender. *The Journal of Econo*mic *Behavior and Organization* 135, 66-81.

Dandy, J., Brewer, N., & Tottman, R. (2001). Self-consciousness and performance decrements within a sporting context. *The Journal of Social Psychology*, *141*(1), 150-152.

De Paola, M., & Scoppa, V. (2015). Gender differences in reaction to psychological pressure: Evidence from tennis players. mimeo.

Doan, B. K., Newton, R. U., Kraemer, W. J., Kwon, Y. H., & Scheet, T. P. (2007). Salivary cortisol, testosterone, and T/C ratio responses during a 36-hole golf competition. *International Journal of Sports Medicine*, 28(6), 470-479.

Dohmen, T. J. (2008). Do professionals choke under pressure? *Journal of Economic Behavior & Organization*, *65*(3), 636-653.

Dohmen, T., & Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *The American Economic Review*, 101(2), 556-590.

Gauriot, R., & Page, L. (2014). Does success breed success? A quasi-experiment on strategic momentum in dynamic contests (No. 028). QUT Business School. mimeo.

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

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *The American Economic Review*, *90*(4), 715-741.

Goleman, D., & Boyatzis, R. (2008). Social intelligence and the biology of leadership. *Harvard Business Review*, *86*(9), 74-81.

González-Díaz, J., Gossner, O., & Rogers, B. W. (2012). Performing best when it matters most: Evidence from professional tennis. *Journal of Economic Behavior & Organization*, *84*(3), 767-781.

Greene, W. (2004). The behavior of the fixed effects estimator in nonlinear models. *The Econometrics Journal*, 7(1), 98–119.

Hickman, D. C., & Metz, N. E. (2015). The impact of pressure on performance: Evidence from the PGA TOUR. *Journal of Economic Behavior & Organization*, *116*(6), 319-330.

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

Jurajda, Š., & Münich, D. (2011). Gender gap in performance under competitive pressure: Admissions to Czech universities. *The American Economic Review*, *101*(3), 514-518 .

Kirschbaum, C., Wüst, S., & Hellhammer, D. (1992). Consistent sex differences in cortisol responses to psychological stress. *Psychosomatic Medicine*, *54*(6), 648-657.

Klaassen, F. J., & Magnus, J. R. (2001). Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model. *Journal of the American Statistical Association*, *96*(454), 500-509.

Kleinjans, K. J. (2009). Do gender differences in preferences for competition matter for occupational expectations? *Journal of Economic Psychology*, *30*(5), 701-710 .

Koning, R. H. (2011). Home advantage in professional tennis. *Journal of Sports Sciences*, *29*(1), 19-27.

Krumer, A., Rosenboim, M., & Shapir, O. M. 2016. Gender, competitiveness, and physical characteristics: Evidence from professional tennis. *Journal of Sports Economics*, *17*(3), 234-259.

Kudielka, B. M., Hellhammer, J., Hellhammer, D. H., Wolf, O. T., Pirke, K. M., Varadi, E., Pilz J., & Kirschbaum, C. (1998). Sex differences in endocrine and psychological responses to psychosocial stress in healthy elderly subjects and the impact of a 2-week dehydroepiandrosterone treatment 1. *The Journal of Clinical Endocrinology & Metabolism*, *83*(5), 1756-1761.

Lautenbach, F., Laborde, S., Achtzehn, S., & Raab, M. (2014). Preliminary evidence of salivary cortisol predicting performance in a controlled setting. *Psychoneuroendocrinology*, *42*, 218-224.

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

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

Macpherson, D. A., & Hirsch, B. T. (1995). Wages and gender composition: why do women's jobs pay less? *Journal of Labor Economics*, *13*(3), 426-471.

Morris, C. (1977). The most important points in tennis. *Optimal Strategies in Sports*, *5*, 131-140.

Neyman, J., & Scott, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrica: Journal of the Econometric Society*, *16*(1), 1-32.

Niederle, M. (2014). Gender (No. 20788). *NBER Working Paper Series*.

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

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

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

Paserman, D. 2010. Gender differences in performance in competitive environments: Evidence from professional tennis players. mimeo.

Polachek, S. W. (1981). Occupational self-selection: A human capital approach to sex differences in occupational structure. *The Review of Economics and Statistics*, *63*(1), 60-69.

Shurchkov, O. (2012). Under pressure: gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, *10*(5), 1189-1213.

Stroud, L. R., Salovey, P., & Epel, E. S. (2002). Sex differences in stress responses: social rejection versus achievement stress. *Biological Psychiatry*, *52*(4), 318-327.

Wooldridge, J. (2002). Econometric Analysis of Cross Section and Panel Data. The MIT Press. Cambridge, Massachusettes, London, England.

Wooldrdge, J. (2006). Introductory Econometrics: a modern approach. Mason, OH, Thomson/South-Western.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit‐formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459-482.

Table 1:
Descriptive statistics

Variable	General Sample: (N=8,280)		Men Only (N=4,153)		Women Only (N=4,127)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Break	0.274	0.446	0.200	0.400	0.349	0.477
PressureBP	0.025	0.010	0.026	0.010	0.024	0.010
PressureGR	0.220	0.138	0.204	0.139	0.236	0.135
Pressure D	0.126	0.332	0.141	0.348	0.111	0.314
Relative round	0.281	0.185	0.282	0.186	0.280	0.184
Clay	0.244	0.430	0.234	0.423	0.254	0.435
Hard	0.492	0.500	0.500	0.500	0.484	0.500
Server's Ranking Index : 8 – $\log_2$ (server's ranking)	2.824	1.854	2.844	1.885	2.804	1.821
Receiver's Ranking Index: 8 - $\log_2$ (receiver's ranking)	2.797	1.845	2.821	1.873	2.774	1.816
GapRank = $\log_2$ (receiver's ranking)-$\log_2$ (server's ranking)	0.027	2.597	0.023	2.614	0.031	2.579
Home advantage of server	0.058	0.234	0.065	0.247	0.051	0.220
Home advantage of receiver	0.057	0.232	0.064	0.244	0.050	0.219
Height of server (cm)	179.671	8.864	185.68	6.506	173.622	6.488
Height of receiver (cm)	179.605	8.859	185.61	6.487	173.557	6.496
BMI of server	22.049	1.843	23.194	1.427	20.896	1.454
BMI of receiver	22.049	1.835	23.186	1.420	20.906	1.454
Game (serial number of game in set)	5.348	2.970	5.485	3.028	5.211	2.906
Women	0.498	0.500				
Matches	886		430		456	

Table 2:
Indexes of pressure

Set status	**Average values of PressureGR**		**PressureBP**
	Women	Men	Women + Men
0-0	0.232	0.116	0.0216
0-1	0.273	0.140	0.0212
0-2	0.123	0.142	0.0177
0-3	0.117	0.066	0.0124
0-4	0.070	0.000	0.0065
0-5	0.000	0.000	0.0018
1-0	0.236	0.213	0.0205
1-1	0.293	0.184	0.0242
1-2	0.246	0.222	0.0238
1-3	0.191	0.123	0.0156
1-4	0.176	0.065	0.0104
1-5	0.006	0.019	0.0052
2-0	0.163	0.224	0.0206
2-1	0.254	0.225	0.0229
2-2	0.298	0.236	0.0267
2-3	0.307	0.199	0.0293
2-4	0.249	0.078	0.0161
2-5	0.087	0.048	0.0089
3-0	0.062	0.067	0.0155
3-1	0.162	0.259	0.0229
3-2	0.259	0.282	0.0299
3-3	0.362	0.286	0.0312
3-4	0.398	0.268	0.0351
3-5	0.202	0.019	0.0144
4-0	0.030	0.011	0.0095
4-1	0.029	0.032	0.0182
4-2	0.114	0.239	0.0266
4-3	0.237	0.259	0.0287
4-4	0.410	0.413	0.0388
4-5	0.536	0.310	0.0447
5-0	0.000	0.000	0.0014
5-1	0.042	0.017	0.0093
5-2	0.062	0.035	0.0100
5-3	0.189	0.253	0.0331
5-4	0.291	0.325	0.0386
5-5	0.451	0.325	0.0437
5-6	0.420	0.328	0.0411
6-5	0.420	0.353	0.0548

Note: The values of PressureBP are taken from Passerman (2010), Table 3.

Table 3
The effect of competitive pressure on the probability to lose a game on serve

	Basic	Basic No controls	Linear game trend	Quadratic game trend	Gender specific game trend	Tie games	Basic – Normalized ranking
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS regression							
PressureGR	0.354***	0.302***	0.335***	0.331***	0.329***	0.295***	0.353***
	(0.051)	(0.049)	(0.052)	(0.052)	(0.053)	(0.093)	(0.051)
PressureGR * Women	-0.148*	-0.137*	-0.136*	-0.145*	-0.128	-0.479***	-0.147*
	(0.081)	(0.079)	(0.082)	(0.083)	(0.083)	(0.156)	(0.081)
Pressure effect among women	0.206***	0.165***	0.199***	0.187***	0.201***	-0.184	0.206***
	(0.064)	(0.062)	(0.064)	(0.066)	(0.064)	(0.125)	(0.064)
Panel B: Logit regression							
PressureGR	0.477***	0.395***	0.458***	0.451***	0.443***	0.669***	0.476***
	(0.071)	(0.067)	(0.071)	(0.071)	(0.071)	(0.193)	(0.071)
PressureGR * Women	-0.309***	-0.250***	- 0.295***	-0.303***	-0.277***	-0.873***	-0.308***
	(0.097)	(0.090)	(0.096)	(0.097)	(0.096)	(0.205)	(0.097)
Pressure effect among women	0.168***	0.144***	0.162***	0.148**	0.166***	-0.204	0.169***
	(0.060)	(0.055)	(0.060)	(0.064)	(0.060)	(0.138)	(0.060)
Game	N	N	Y	Y	Y	N	N
$Game^2$	N	N	N	Y	N	N	N
Game * Women	N	N	N	N	Y	N	N
Match fixed effects	Y	Y	Y	Y	Y	Y	Y
Observations	8,280	8,280	8,280	8,280	8,280	2,552	8,280
Number of Matches	886	886	886	886	886	886	886

*** p<0.01, ** p<0.05, * p<0.1

All regressions include match fixed effects. Standard errors clustered at the match level are in parentheses. Additional control variables include the difference between the server's and the receiver's rankings, a dummy for the home advantage of the server, the height of the server and the BMI of the server.

Table 4:
The effect of competitive pressure on the probability to lose a game on serve

	Entire Sample				Discontinuity sample	
	Basic	Linear game trend	Quadratic game trend	Gender specific game trend	Basic	Linear game trend
	(1)	(2)	(3)	(4)	(5)	(6)
<u>Panel A: OLS regression</u>						
PressureGR	0.648***	0.638***	0.635***	0.639***	0.725***	0.639***
	(0.062)	(0.064)	(0.064)	(0.065)	(0.114)	(0.119)
PressureGR * Women	-0.301***	-0.294***	-0.302***	-0.296***	-0.366**	-0.373**
	(0.088)	(0.089)	(0.090)	(0.091)	(0.159)	(0.161)
Pressure effect among women	0.347***	0.344***	0.333***	0.343***	0.359***	0.266**
	(0.063)	(0.063)	(0.065)	(0.063)	(0.112)	(0.114)
<u>Panel B: Logit regression</u>						
PressureGR	1.115***	1.099***	1.098***	1.086***	1.354***	1.178***
	(0.103)	(0.103)	(0.104)	(0.104)	(0.158)	(0.167)
PressureGR * Women	-0.736***	-0.736***	-0.736***	-0.706***	-0.820***	-0.867***
	(0.130)	(0.130)	(0.130)	(0.132)	(0.224)	(0.221)
Pressure effect among women	0.379***	0.364***	0.362***	0.380***	0.534***	0.311
	(0.076)	(0.077)	(0.079)	(0.078)	(0.201)	(0.203)
Game	N	Y	Y	Y	N	Y
$Game^2$	N	N	Y	N	N	N
Game * Women	N	N	N	Y	N	N
Match fixed effects	Y	Y	Y	Y	Y	Y
Match fixed effects * OG	Y	Y	Y	Y	Y	Y
Observations	8,280	8,280	8,280	8,280	3,850	3,850
Number of matches	886	886	886	886	886	886
*** p<0.01						

All regressions include match fixed effects and interactions between match fixed effects and a dummy indicating an odd game. Standard errors clustered at the match level are in parentheses. No control variables are included.

Table 5:
The effect of competitive pressure on the probability to lose a game on serve by gender

	All Games				Tie games	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
PressureGR	0.348***	0.211***	0.648***	0.347***	0.295***	-0.184
	(0.051)	(0.064)	(0.062)	(0.063)	(0.093)	(0.126)
Match fixed effects	Y	Y	Y	Y	Y	Y
Match fixed effects * OG	N	N	Y	Y	N	N
Observations	4,153	4,127	4,153	4,127	1,363	1,189
Number of matches	430	456	430	456	430	456

*** $p<0.01$

All regressions include match fixed effects. Standard errors clustered at the match level are in parentheses. Additional control variables include the difference between the server's and the receiver's rankings, a dummy for home advantage of the server, the height of the server and the BMI of the server.

Table 6:
Robustness analysis

	IV estimates - PressureGR				Pressure D	
	Men	Women	Men	Women	All Data	
	(1)	(2)	(3)	(4)	(5)	(6)
Pressure	0.554***	0.126	1.134***	0.462***	0.072***	0.079***
	(0.110)	(0.100)	(0.124)	(0.106)	(0.023)	(0.023)
Pressure * Women					-0.080**	-0.080**
					(0.032)	(0.032)
Linear game trend	Y	Y	Y	Y	Y	N
Match Fixed Effects	Y	Y	Y	Y	Y	Y
Match Fixed Effects * OG	N	N	Y	Y	N	Y
Observations	4,153	4,127	4,153	4,127	8,280	8,280
Number of matches	430	456	430	456	886	1,772

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

All regressions include match fixed effects. Standard errors clustered at the match level are in parentheses. Additional control variables include difference between the server's and the receiver's rankings, a dummy for home advantage of server, the height of the server and the BMI of the server. In all columns PressureBP is used as an instrumental variable.

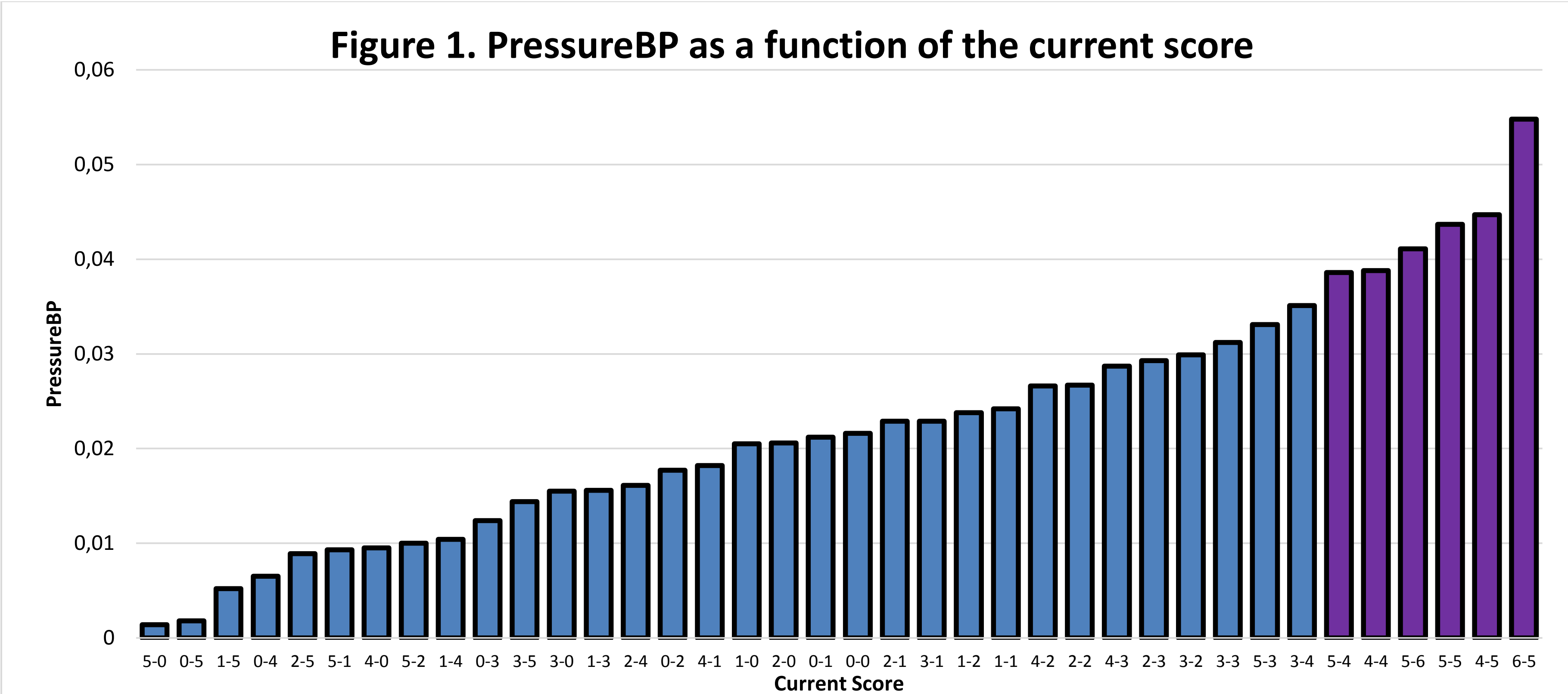
Figure 1. PressureBP as a function of the current score
0,06
0,05
0,04
0,03
0,02
0,01
0
PressureBP
5-0 0-5 1-5 0-4 2-5 5-1 4-0 5-2 1-4 0-3 3-5 3-0 1-3 2-4 0-2 4-1 1-0 2-0 0-1 0-0 2-1 3-1 1-2 1-1 4-2 2-2 4-3 2-3 3-2 3-3 5-3 3-4 5-4 4-4 5-6 5-5 4-5 6-5
Current Score

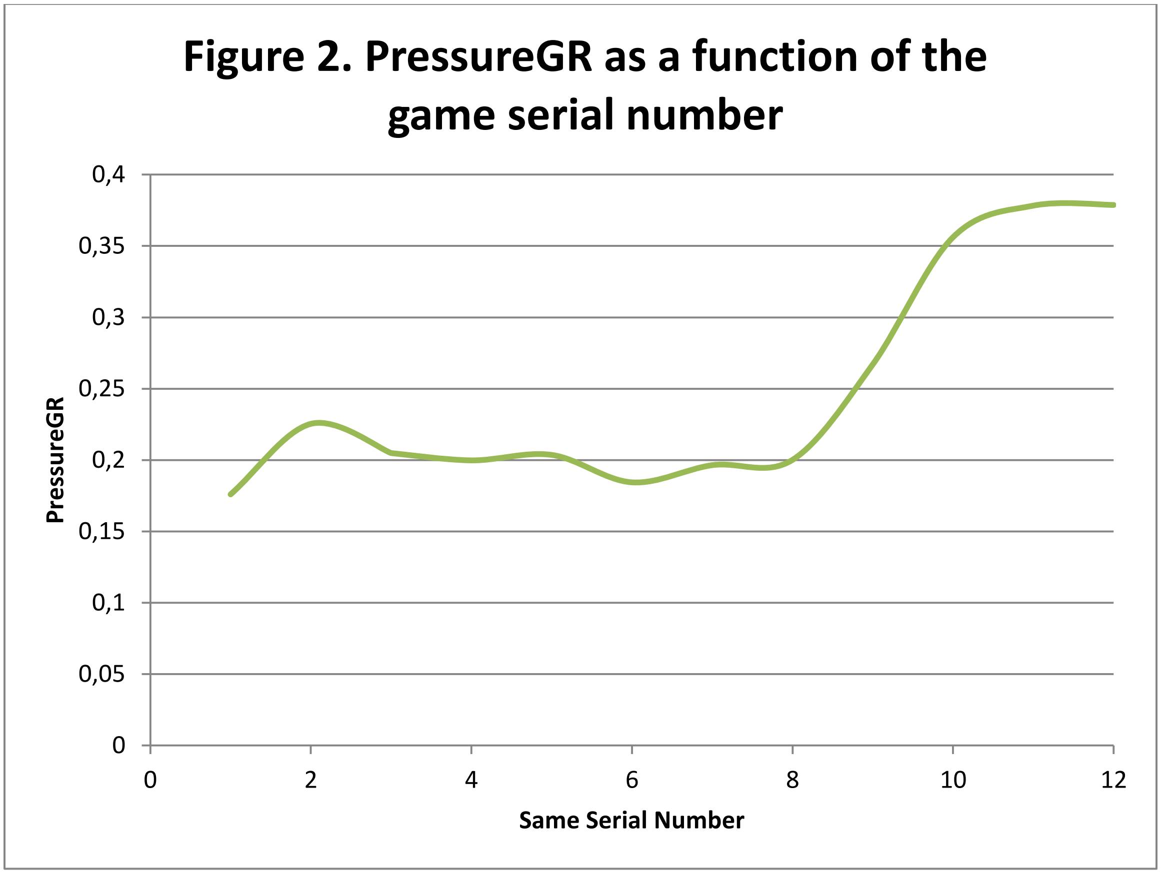
Figure 2. PressureGR as a function of the game serial number
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
PressureGR
0
2
4
6
8
10
12
Same Serial Number

Appendix

Table A1:
List of Sources

www.atpworldtour.com
www.tennisbetsite.com
www.tennisexplorer.com
2010.australianopen.com
2010.wimbledon.org
2010.rolandgarros.com
2010.usopen.org
www.tennisnewsonline.com
eurosport.du.ae
www.dailymail.co.uk

For more information about the links and the sites, the authors can be contacted by e-mail.

Table A2:
Selection Bias

	Men			Women		
	Omitted data	Included data	Difference	Omitted data	Included data	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Average ranking	2.703	2.833	0.130	2.777	2.794	0.016
	(1.309)	(1.355)	(0.167)	(1.253)	(1.271)	(0.184)
Abs gap ranking	1.984	2.133	0.149	2.314	2.058	-0.257
	(1.392)	(1.562)	(0.190)	(1.830)	(1.603)	(0.236)
Average BMI	23.142	23.179	0.037	20.704	20.905	0.201
	(0.979)	(0.987)	(0.125)	(1.109)	(1.000)	(0.152)
Average height	185.007	185.641	0.634	174.592	173.568	-1.024
	(4.447)	(4.767)	(0.584)	(4.187)	(4.663)	(0.695)

Columns 1 and 4 present the average value of each of the characteristics for the omitted matches for men and women, respectively. Similarly, Columns 2 and 5 refer to the included matches for men and women, respectively. Standard deviations are in parentheses. Results from univariate regressions of each of the variables in this table on a dummy variable indicating whether the observation belongs to the included or omitted dataset appear in Columns 3 and 6, for men and women, respectively. Standard errors are in parentheses.

Table A3:
GLS estimates of the effect of competitive pressure on the probability to lose a game on serve

	Basic	Basic No controls	Linear game trend	Quadratic game trend	Gender specific game trend	Tie games
	(1)	(2)	(3)	(4)	(5)	(6)
PressureGR	0.556***	0.258***	0.539***	0.537***	0.526***	0.725***
	(0.079)	(0.069)	(0.080)	(0.080)	(0.080)	(0.278)
PressureGR * Women	-0.331***	-0.131	-0.317***	-0.322***	-0.302***	-0.978***
	(0.112)	(0.100)	(0.111)	(0.113)	(0.112)	(0.368)
Pressure effect among women	0.225***	0.127*	0.221***	0.215***	0.224***	-0.254
	(0.079)	(0.073)	(0.078)	(0.084)	(0.079)	(0.242)
Game	N	N	Y	Y	Y	N
$Game^2$	N	N	N	Y	N	N
Game * Women	N	N	N	N	Y	N
Match fixed effects	Y	Y	Y	Y	Y	Y
Observations	8,280	8,280	8,280	8,280	8,280	2,552
Number of Matches	886	886	886	886	886	886

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

All regressions include match fixed effects. Standard errors clustered at the match level are in parentheses. Additional control variables include the difference between the server's and receiver's rankings, dummy for home advantage of the server, height of the server and BMI of server.