Gender Differences in Performance in Competitive Environments? Evidence from Professional Tennis Players*

M. Daniele Paserman
Boston University and Hebrew University
paserman@bu.edu

January 2010

Abstract

This paper uses detailed stroke-by-stroke data from seven tennis Grand Slam tournaments played between 2006 and 2007 to assess whether men and women respond differently to competitive pressure in a setting with large monetary rewards. It finds that at crucial junctures of the match, both men and women adopt a more conservative and less aggressive playing strategy, meaning that the probability of hitting winning shots and making unforced errors decreases. The odds of making an unforced error relative to hitting a winner fall for women, while they remain constant for men. However, using a simple game-theoretic model, I argue that the men's game deteriorates at least as much as the women's game on more important points. I estimate that, for both men and women, the probability that a player wins a match against an opponent of equal quality would increase from 0.5 to about 0.75-0.80 if he or she could avoid the deterioration in performance on more important points.

JEL codes: J16, J24, J71, L83, M50.
Keywords: Gender differences, performance under pressure, tennis.

* This paper is a substantially revised version of "Gender Differences in Performance in Competitive Environments: Evidence from Professional Tennis Players," CEPR Discussion Paper 6335, June 2007. I thank Dan Ariely, Alison Beach, Iván Fernández-Val, Joshua Lustig, Franc Klaassen, Esteban Klor, Lawrence Katz, Motty Perry, Zhongjun Qu, Antonio Rangel, Avi Simhon, Eyal Winter, seminar participants at Hebrew University, the University of Pittsburgh, UC Davis, University of Illinois at Urbana-Champaign, Boston University, Case Western Reserve University, Tufts University, Tel Aviv University, Yale University, Brown University, Harvard University, Boston College and conference participants at the European Summer Symposium in Labor Economics for comments and suggestions. This paper would not have been possible without Yaron Aronshtam's outstanding research assistance. Excellent research assistance was also provided by Shinsuke Tanaka and Jordi Vidàl. All errors are my own.

Recent decades have seen a dramatic increase in female labor force participation rates, and a considerable narrowing of the gender gap. Yet, despite these advances, the gender gap in wages is still substantial, even if one looks at men and women with the same amount of experience and education. At the very high end of the wage distribution, women have found it particularly hard to break the glass ceiling and to make inroads into the upper echelons of management, academia and prestigious professions. Bertrand and Hallock (2001) report that among the highest paid executives at top corporate firms between 1992 and 1997, only 2.5 percent are women.[1]

Several explanations have been put forth for this phenomenon, ranging from discrimination to differences in preferences, each of which can then lead to differential investments in human capital and on-the-job training. In a recent paper, Gneezy, Niederle and Rustichini (2003) have put forward an intriguing hypothesis for the large under-representation of women in high-powered jobs: women may be less effective than men in competitive environments, even if they are able to perform similarly in a non-competitive environment. Using experimental evidence, they revealed the existence of a significant gender gap in performance in a tournament setting where wages were based on a winner-takes-all principle, while no such gap existed when players were paid according to a piece rate. The reason for this gap is that men's performance increases significantly with the competitiveness of the environment, while women's performance does not. Similar findings were also obtained by Gneezy and Rustichini (2004), who analyzed the performance of young boys and girls in a race over a short distance. Niederle and Vesterlund (2007), on the other hand, found no gender differences in performance on an arithmetic task under either a non-competitive piece rate

[1] A similar underrepresentation of women is also found among CEOs at Fortune 500 companies (http://money.cnn.com/magazines/fortune/fortune500/womenceos), tenured faculty at leading research institutions (MIT, 1999), conductors of philharmonic orchestras in the U.S. (http://www.infoplease.com/ipea/A0106174.html,) , or top surgeons in New York City according to *New York* magazine (http://nymag.com/bestdoctors/).

compensation scheme, or a competitive tournament scheme. However, they found that men were significantly more likely than women to select into the competitive compensation scheme when given the choice, and that such a choice could not be explained by performance either before or after the entry decision. A similar behavior was also found by Booth and Nolen (2009) and Dohmen and Falk (2006), who attributed part of the gender difference in preferences for the competitive environment to differences in the degree of risk aversion.

Some recent studies have analyzed the link between performance and competition in non-experimental settings. Örs et al. (2008) find that performance is statistically lower for women in a high-stakes entrance exam for a highly competitive business school in France. Similarly, Price (2008) finds that male graduate students made quicker progress in their doctoral studies in response to the institution of a highly competitive fellowship program, while there was no response by women. Lavy (2008a) found that the gender gap in test scores (which favors girls) on "blind" high school matriculation exams (i.e., exams that are graded by an external committee) is smaller than the gender gap in scores on exams graded internally. In contrast, Goldin and Rouse (2000) found that female orchestra musicians were more likely to advance through the stages of the hiring process if gender-blind auditions were in place. Both papers attribute these findings to discrimination, but one cannot rule out that women underperform when they face increased pressure, (which can take the form of the blind exam for high-school students, or performance in front of a predominantly male selection committee in the case of musicians).[2]

[2] On the other hand, Lavy (2008b) finds no gender differences in performance in a tournament in which contestants have more time to prepare and plan their strategies, and Manning and Saidi (2008) argue that gender differences in the incidence of pay-for-performance schemes can account for only a small fraction of the gender gap in the United Kingdom.

There is also a vast literature in social psychology on the phenomenon of "choking under pressure," i.e., suboptimal performance despite a high degree of achievement motivation (Baumeister, 1984).[3] Recently, Ariely et al. (2005) find evidence that high reward levels can have detrimental effects on performance. When they test specifically for gender differences in performance under pressure, however, they do not find any evidence that women do relatively worse when they have to perform a task while being observed by others. Dohmen (2006) presents field evidence from male professional soccer players, and finds that performance is not affected by the degree of competitive pressure, but is negatively affected by the presence of a supportive audience.

In this paper, I complement the existing literature by examining whether men and women respond differently to competitive pressure in a setting with large monetary rewards. Specifically, I investigate whether players exhibit a deterioration in performance at the crucial stages of the match. I focus on professional tennis players in seven Grand Slam tournaments played between 2006 and 2007. One of the advantages of using sports data is that it is possible to observe detailed measures of productivity or performance. In particular, I use information on the length rallies and on the way the rally ends to construct a *stroke-by-stroke* data set and infer how the performance of players changes over the course of a match. This is achieved by setting up a structural econometric model that estimates the probability hitting winning shots and unforced errors at every stroke.

For tennis, a natural measure of performance would appear to be the incidence of *unforced errors*, which, by definition, are errors that cannot be attributed to any factor other than

[3] Previous research has highlighted a number of possible sources of pressure that may induce decreased performance: competitive conditions, or the magnitude of the stakes or rewards to be achieved (Baumeister, 1984), the importance of achieving a success (Kleine, Sampedro, and Lopes, 1988), or the expectations of outside observers (Baumeister et al., 1985). The presence of a supportive audience (Butler and Baumeister, 1998), or the mere presence of others might also create pressure and induce individuals to choke (Zajonc, 1965).

poor judgment and execution by the player. However, I show using a simple model that the incidence of unforced errors is not necessarily informative about the quality of one's play. In fact, it is entirely possible that the probability of winning a point increases even as the share of strokes that are unforced errors increases. This apparently counterintuitive statement has a quite simple explanation: what matters for the probability of winning the point is the *absolute* probability of hitting winning shots and unforced errors. If a relative increase in the ratio of unforced errors to winning shots is accompanied by an increase in the absolute gap between winning shots and unforced errors, the probability of winning the point will increase. Therefore, the incidence of unforced errors on its own is not an appropriate indicator of performance. Instead, I will use as my preferred performance metric the probability of winning the point against an "average" opponent, given the estimated probabilities of hitting winning shots and unforced errors.

One of the key features of the analysis is the construction of a refined measure of the decisiveness of each point. Following Klaassen and Magnus (2001), I define the *importance* of each point as the difference in the probability of winning the entire match as a result of winning or losing the current point. To calculate the importance of each point, I use a dynamic programming algorithm that takes into account the structure of a match in a Grand Slam tournament, and calculates recursively, for every pairing of players, the probability of winning the match at every possible stage. One of the remarkable features of the importance variable is that it evolves very non-linearly over the course of a match, generating an abundance of useful variation that can be used for estimation.

I find that there are important differences in the way the game is played as points become more important. Both men and women hit fewer winning shots and fewer unforced errors on more important points. The ratio of unforced errors to winning shots increases with importance for women, but remains roughly constant for men. This implies that at crucial stages of the match women's points, but not men's, are more likely to end in unforced errors. All these results are robust to the sample used, to the exact construction of the importance variable, and to the specification of the econometric model. The relative increase in the odds of unforced errors for women, however, does not signify a relative decrease in performance. In fact, given the implied probabilities of hitting winning shots and unforced errors, the probability of winning the point against a hypothetical "average" player decreases by about the same amount for both men and women. This deterioration in performance is substantial: I estimate that in a match between two players of equal ability, the player that is able to prevent the deterioration in performance on more important points can increase the probability of winning the match from 50 percent to 75-80 percent. This is equivalent to a more than fourfold increase in player ranking.

The results imply that top professionals play a more conservative and less aggressive strategy as points become more important. Playing a safe strategy implies a lower probability of hitting outright winners as well as unforced errors. What then can explain this behavior? One may be inclined to attribute these differences to risk attitudes (Croson and Gneezy, 2004). This explanation should be promptly dismissed. Tennis is a zero-sum game, and as such, there is no tradeoff between mean and variance. If a player has to choose between playing an aggressive stroke that will make her win the point with a 55% probability and make an error with a 45% probability, or playing a safe stroke which will prolong the rally and result in a 50% probability of winning the point, it is clearly optimal to choose the former.

Instead, I argue using a simple game-theoretic model that a shift from an aggressive to a less aggressive strategy may arise as an optimal response to a change in the intrinsic probabilities of hitting winning shots or unforced errors. Simply put, if players know that on more important points they are more likely to make unforced errors when playing aggressively (or, conversely, less likely to hit winning shots), they will revert to a safer playing strategy.

While these empirical findings are certainly intriguing, one should keep in mind that the results presented here are relevant in this specific context, and one should exercise caution before extending the conclusions to athletes in other sports, let alone to managers, surgeons, or other professionals who must make quick and accurate decisions in high pressure situations. Nevertheless, the fact that we find such robust evidence of underperformance in high-stakes situations, even among experienced professional drawn from the extreme right tail of the talent distribution, is without a doubt a novel finding that raises some interesting questions about the effect of high-powered incentive contracts, and should stimulate further investigation.

The rest of the paper is structured as follows. The next section introduces some basic terminology and concepts in the game of tennis. Section 2 presents a simple non-strategic model of tennis that clarifies the relationship between performance and observable variables, and shows how data on the number of strokes in the rally and on the way the rally ends allows to identify the underlying probabilities of hitting winners and unforced errors. Section 3 presents a strategic model of tennis, which shows how the underlying probabilities of hitting winning shots and unforced errors determine whether players choose to play aggressive or safe strategies in equilibrium. Section 4 introduces the data, and describes the construction of the importance variable and the econometric methodology. Section 5 presents the basic results of the analysis, estimates the degree to which performance deteriorates on more important points, and assesses the robustness of the results to alternative specifications of the sample, the importance variable and the econometric methodology. Section 6 presents a simulation that illustrates how differences in performance at crucial stages of the game can have quite dramatic effects on the overall probability of winning the match. Section 7 discusses the results and concludes.[4]

1. Tennis: basic concepts

Rules. Tennis is a game played by two players who stand on opposite sides of a net and strike a ball in turns with a stringed racket. Their objective is to score *points* by striking the ball within a delimited field of play (the *court*) and out of the reach of the opponent. The scoring system in tennis is highly non-linear. A tennis *match* comprises an odd number of *sets* (three or five). A *set* consists of a number of *games* (a sequence of points played with the same player serving), and games, in turn, consist of *points*. The match winner is the player who wins more than half of the sets. Typically, a player wins a set when he wins at least six games and at least two games more than his opponent.[5] A game is won by the first player to have won at least four points and at least two points more than his opponent.

Typology of points. A point is lost when one of the players fails to make a legal return of the ball. This can happen in a number of ways: a *winner* is a forcing shot that cannot be reached by the opponent and wins the point; a *forced error* is an error in a return shot that was forced by the opponent; an *unforced error* is an error in a service or return shot that cannot be attributed to any factor other than poor judgment and execution by the player. The definition of unforced

---

[4] In a previous version of the paper (Paserman, 2007), I had found a significant gender difference in the effect of importance on the proportion of points ending in unforced errors, and I had concluded that this implied that women were more likely to exhibit a deterioration in performance in high pressure situations. The result that women's points are more likely to end in unforced errors at crucial stages of the match is confirmed in this version of the paper, but the introduction of carefully collected data on rally length allows me estimate the probability of hitting unforced errors and winning shots at every *stroke*. This was not feasible with the version of the data set used in the previous version. I am therefore able to assess more accurately the effect of importance on player performance by specifying a full structural model for tennis points. This version also differs from the previous one in that it analyzes a substantially larger data set (seven Grand Slam tournaments instead of four).

[5] If the score in games is tied at 6-6, players usually play "tie-break", which is won by the player who first reaches seven points, with a margin of at least two points over his opponent.

errors is critical for the purposes of this paper. Statistician Leo Levin of IDS Sports, who has compiled statistics for all the major tennis tournaments, argues that the idea behind unforced errors is to place the blame for an error on one of the two players. He defines an unforced error as a situation when a player has time to prepare and position himself to get the ball back in play and makes an error.[6] In practice, the classification of points into the three categories is made by courtside statistics-keepers (usually amateur tennis players with a substantial amount of experience in both playing and watching tennis matches) who are recruited and trained specifically in advance of the tournament.[7]

2. A Non-strategic Model of Tennis Points

I present here a simple model of tennis points that will help to clarify the relationship between the quantities observed in the data and performance. I begin by describing a simple non-strategic model, where the outcome of each point is mechanically determined by the exogenous probabilities of hitting winning shots and unforced errors. The model helps us to understand whether changes in the outcomes of tennis points can be interpreted as changes in the underlying performance parameters. I then present a strategic model in which the equilibrium outcome of each point is determined by the players' strategies and the underlying performance parameters.

The model

---

[6] From *http://www.tennis.com/yourgame/asktheeditors/asktheeditors.aspx?id=1432*, last viewed on April 26, 2007.

[7] I do not have information on the identity of the statistics keepers, so a priori one cannot rule out the possibility that any observed gender difference in performance are simply the result of gender bias on their part. Note however that the argument in this paper is based not only on the average level of performance, but on how performance varies with the importance of the point. Therefore, for our results to be just an artifact of the statistics keepers' perceptions, it would be necessary that they are more likely to classify women's errors as unforced only when these errors occur at crucial stages of the match.

As discussed above, every tennis point must end in one of three possible ways: an outright winner, an unforced error, or a forced error. Therefore, each *stroke* must be classified in one out of five mutually exclusive categories: a) an unforced error; b) a winner; c) a forcing shot that induces a forced error; d) a forced error; and e) a ball that is put in play and that the opponent is able to return. For all intents and purposes, a forcing shot that induces a forced error can be treated as equivalent to a winning shot: it is a shot that allows the player to win the point. Therefore, I collapse winning forcing shots with outright winners, and ignore forced errors.

The probability of hitting a winning shot or an unforced error is determined mechanically by the players' ability, and does not depend on the history of the rally up to that point.[8] (this assumption can be relaxed). Specifically, let w_A and w_B be the probabilities of hitting a winning shot, and let u_A and u_B be the probabilities of hitting an unforced error, for players A and B, respectively. Then, the probabilities of putting the ball in play are p_A and p_B respectively. Assume that player A strikes the ball first. The probability that player A wins the point is:

$$P(\text{player A wins point}) \equiv q_A = w_A + p_A u_B + p_A p_B w_A + p_A^2 p_B u_B + \ldots$$
$$= \frac{w_A + p_A u_B}{1 - p_A p_B}.$$

It follows immediately that the probability that player A wins the point is increasing in w_A ($\partial q_A / \partial w_A > 0$) and decreasing in u_A ($\partial q_A / \partial u_A < 0$). Holding everything else constant, an increase in the probability of hitting winning shots raises the probability of winning the point, and increase in the probability of hitting unforced errors lowers the probability of winning the point.

More interesting is the analysis of the effect of a change in w_A and u_A such that the ratio u_A/w_A remains constant. The following proposition holds:

[8] This assumption will be relaxed in the strategic model. I also abstract from the special role of serves in tennis, even though the empirical analysis will take serves into account.

Proposition 1: If $w_A = \alpha \tilde{w}_A$ and $u_A = \alpha \tilde{u}_A$ ($\alpha > 0$), then $sign\left(\frac{\partial q_A}{\partial \alpha}\right) = sign(\tilde{w}_A w_B - \tilde{u}_A u_B)$.

Proof: Follows from simple differentiation of q_A.

The implication of this proposition is that if winners are relatively frequent, then an increase in α (i.e., an equiproportionate increase in winners and unforced errors) *raises* the probability that player A wins the point. The intuition for this result is straightforward and can be illustrated by a simple example. Consider an extreme case where both players have $w_A = w_B = 0.1$ and $u_A = u_B = 0.0001$. That is, both players hardly ever make mistakes. If both w_A and u_A double, clearly player A is more likely to win the point, because unforced errors are so rare that they essentially never happen. By continuity of q_A, it is clear that the probability that player A wins the point may increase even if u_A/w_A increases (i.e., player A hits relatively more unforced errors).

The key message to take from the proposition is that to assess player performance, we need information on the absolute values of w_j and u_j, and not just u_j/w_j. In what follows, I show that data on rally length and on the way in which the rally ends is sufficient to identify these key probabilities.

Identification

To illustrate the key to identification, assume for simplicity that players have identical ability: $w_A = w_B = w$; $u_A = u_B = u$. Then, the probability that a point ends in unforced error (by either player) is $P_u = \frac{u}{u+w}$. Similarly, the probability that a points ends in a winning shot is $P_w = \frac{w}{u+w}$. Since $P_u + P_w = 1$, these two equations are not independent, and the underlying parameters *u* and *w* are not identified. In other words, information on how points end alone is not sufficient to identify the performance parameters.

However, assume that we also have information on rally length. The expected number of strokes in a rally is equal to $\frac{1}{u+w}$. It is easy to see that information on the number of strokes in the rally <u>and</u> on how points end is sufficient to identify *u* and *w*.

The same basic ideas apply in a model with heterogeneous abilities. With only information on how the rally ends, we observe four moments. For example, the probability that the rally ends in an unforced error by player A is $\frac{w_A}{1-p_A p_B}$, and the other probabilities are defined analogously. Because only three of these moments are independent we are unable to identify the four structural parameters. However, if we also observe rally length, we can use the fact that the expected number of strokes is $\frac{1+p_A}{1-p_A p_B}$, and we therefore have an additional moment that enables us to identify all the parameters.

3. A Strategic Model of Tennis Points

Of course, it is somewhat unappealing to assume that the outcome of a tennis point is determined completely mechanically, and does not depend on the strategic interactions between the two players. I present here a strategic model of tennis, which will help in explaining the observed outcomes. In particular, the model shows that if at different points in the game players choose different equilibrium strategies, it must be the case that the intrinsic probabilities of hitting winners and unforced errors have changed.

I model the strategic interaction for a single tennis point between two players, A and B, whose objective is to maximize the probability of winning the point. I abstract here from dynamic considerations, whereby winning one point may affect the probabilities of winning subsequent points in the match. Player 1 makes the first move (hits the first stroke of the rally),[9] and can choose one of two actions, to play "soft" (action 0) or to play "aggressive" (action 1). Player A's stroke will result in any one of three outcomes: 1) the player hits a winner (or induces a forced error by the opponent) and wins the point; 2) the player hits an unforced error and loses the point; and 3) the ball is put in play, and the opponent has the chance to hit a stroke back. The opponent (Player B) in turn chooses between playing softly or aggressively, and his/her stroke results in either a winner, an unforced error, or the ball returning to Player A. The two players continue exchanging strokes sequentially until one of the two players hits a winner or commits an unforced error.

The probability of hitting a winner or an unforced error is a function of the player's aggressiveness. Specifically, let w_{0j} be the probability that player *j (j* = A, B) hits a winner when he/she plays soft, and let w_{1j} be the probability of hitting a winner when playing aggressively. Similarly, let u_{0j} be the probability that player *j (j* = A, B) makes an unforced error when he/she plays soft, and let u_{1j} be the probability of making an unforced error when playing aggressively. I also define $p_{0j} = 1 - u_{0j} - w_{0j}$ and $p_{1j} = 1 - u_{1j} - w_{1j}$ as the probabilities of hitting the ball in play and allowing the opponent to make a stroke. I make the following assumption about the basic probabilities:

Assumption 1. a) $w_{1j} > w_{0j}$, for j = A, B **;** b) $u_{1j} > u_{0j}$, for j = A, B **;** c) w_{1j}, w_{0j}, u_{1j}, and u_{0j} are constant over the course of the rally.

[9] For simplicity I abstract from service shots, which involve a different dynamic.

Part a) states that the probability of hitting a winner is higher when playing aggressively than when playing softly – this is almost by definition the meaning of playing aggressively. Part b) states that the probability of making unforced errors is also higher when playing aggressively. This too is an extremely natural assumption: when playing more aggressively, there is a higher risk of losing control of the ball and making mistakes. The final assumption states that the probabilities of hitting winners and unforced errors do not depend on the history of the rally up to that point. In other words, I rule out fatigue effects and strategic buildups of points (e.g., making the opponent run from one side of the court to the other, in order to increase the chances of hitting a winner on the following stroke). While this assumption is probably unrealistic, it greatly simplifies the model, allowing me to concentrate on the salient aspects of the game. In the last part of this section I will discuss some extensions to the model, including allowing the probabilities to depend on the stroke played previously.

Assumption c) implies that the game is stationary: the decision problem of Player A on the third stroke of the rally is exactly identical to the decision problem on the first stroke of the rally. Therefore, I will restrict attention only to *stationary* strategies, i.e., strategies in which the player chooses the same action every time it is his or her turn to strike the ball. I will also restrict attention to subgame perfect equilibria, in which players never choose a suboptimal action off the equilibrium path.

I denote $V^j(\sigma_A, \sigma_B)$ the value to player *j* when player A chooses σ_A at every stroke, and player B chooses action σ_B at every stroke. Denote instead by $U^A(s, \sigma_B)$ the value to player A of choosing *s* the first time it is his or her turn to strike the ball, and then choosing the optimal stationary strategy at every subsequent stroke, when player B chooses action σ_B at every stroke.

$U^B(\sigma_A, s)$ is defined analogously. We can write down these value functions explicitly. For example,

$$U^A(0,0) = w_{0A} + p_{0A}u_{0B} + p_{0A}p_{0B} \max\{V^A(0,0), V^A(1,0)\}.$$

The above expression gives us the probability that player A wins the point when he or she chooses to play Soft on the first stroke, and then plays optimally at every subsequent stroke, when player B also plays Soft. With probability w_{0A} player A hits a winner and wins the point on the first stroke; with probability p_{0A} the ball reaches player B, who then commits an unforced error with probability u_{0B} – hence the probability of winning the point on the second stroke is $p_{0A}u_{0B}$; finally, with probability $p_{0A}p_{0B}$ the ball lands back to player A, who then chooses the optimal strategy between playing softly and aggressively, conditional on player B playing softly at every stroke. In any subgame perfect equilibrium, it must be that the action chosen by player A on the first stroke satisfies:

$$\hat{s}^A(\sigma_B) = \arg\max\{U^A(0,\sigma_B), U^A(1,\sigma_B)\},$$

and the action chosen by player B satisfies:

$$\hat{s}^B(\sigma_A) = \arg\max\{U^B(\sigma_A,0), U^B(\sigma_A,1)\}.$$

To further simplify the model, I make the additional assumption that the two players have identical abilities,[10] hence $w_{1A} = w_{1B} = w_1$; $w_{0A} = w_{0B} = w_0$; $u_{1A} = u_{1B} = u_1$; $u_{0A} = u_{0B} = u_0$.[11] One can now state Proposition 2.

[10] This assumption will also be relaxed later.
[11] The main result of the Proposition, that multiple subgame perfect equilibria cannot arise, holds even in the case in which players have different abilities. Details available upon request.

Proposition 2: In the game described above, when players have identical abilities, only one of three outcomes is possible: a) a unique pure-strategy subgame perfect equilibrium in which both players play "soft"; b) a unique pure-strategy subgame perfect equilibrium in which both players play "aggressive"; c) a degenerate situation in which both players are always indifferent between playing "soft" and "aggressive", and therefore an infinite number of mixed-strategy subgame perfect equilibria may arise.

Proof: See Appendix.

The implication of Proposition 2 is clear: if different types of equilibria arise at different points of the match, it must be the case that the intrinsic probabilities of hitting winners or unforced errors (i.e., the basic parameters of the model – u_0, u_1, w_0, and w_1) have changed. The proof of the proposition shows that the (Soft, Soft) equilibrium will arise if the following condition holds:

$$\textbf{C1: } u_1 > \frac{u_0^2 + w_0\left(w_1 - w_0\right) + \left(1 - u_0 - w_0\right) w_1 u_0}{u_0 + \left(1 - u_0 - w_0\right) w_0}.$$

The inequality reflects the straightforward intuition that if the probability of committing unforced errors when playing aggressively rises, players are more likely to gravitate towards the (Soft, Soft) equilibrium. This same inequality can also be rewritten in terms of w_1, in which case we obtain that when the probability of hitting winners rises when playing aggressively, players will gravitate towards the (Aggressive, Aggressive) equilibrium.

<u>*Extensions*</u>

Players of unequal ability. Allowing players to have different abilities does not alter the basic result of the model. With heterogeneous abilities, it is possible to obtain asymmetric equilibria, where one of the players chooses Soft and the other chooses Aggressive. What is important, though, is that even in this case there can be no multiple equilibria, unless we have a knife-edge configuration of the parameters. Again, a shift in equilibrium strategies must imply that the intrinsic probabilities of winners and unforced errors must have changed. See the Appendix for a modified version of Propositions 1 and 2 that allow for heterogeneous abilities.

Non-stationarity. There are many instances in tennis in which a player has no choice but to play a defensive shot just to keep the rally alive. Therefore, one may question the validity of assumption (1c), that the intrinsic probabilities remain constant over the course of the rally. We can allow defensive shots with a slight modification of the model. Assume that when faced with a hittable shot, players can still choose between two strategies, Soft and Aggressive. However, each stroke entails now four possible outcomes: a winner (with probability w), an unforced error (with probability u), a forcing shot (with probability f), and a hittable shot (with probability $1-w-u-f$). A forcing shot is a stroke that constrains the opponent to play a defensive strategy. Specifically, after a forcing shot, the opponent can either make a (forced) error, with probability e, or return a hittable shot (with probability $1-e$). It turns out that the basic point of Proposition 2 holds even under this modified model: multiple equilibria are possible only under knife-edge values of the parameters. Hence, changes in equilibrium strategies over the course of a match must imply that the intrinsic probabilities have changed.

Numerical examples

Proposition 2 shows that different intrinsic probabilities of unforced errors and winners will result in different playing strategies and in different outcomes of the rally. Table 1 illustrates these different outcomes for selected parameter values. The parameter values were chosen to roughly calibrate the percentage of points ending in unforced errors and the number of strokes per rally in women's matches. The first two rows describe the outcome of a point between two players of identical ability. The only difference between the two rows is that in the first row the players are slightly more likely to hit winners when playing Aggressive ($w_1 = 0.12$ instead of 0.10). This small difference translates into a different equilibrium, a substantially higher probability that the point ends in an unforced error, and in a nearly 20% increase in the length of the rally. The third row presents the equilibrium strategies and outcomes when one of the players has $w_1 = 0.12$, and the other player has $w_1 = 0.10$. Under these probabilities, the equilibrium is still (Aggressive, Aggressive), while the probability of the rally ending in an unforced error and the expected length of the rally lie somewhere in between the outcomes observed in rows 1 and 2. Note that the probability that the better player wins the rally rises by more than two percentage points. We will see later that this can translate into substantially higher probabilities of winning the entire match.

4. Data and Methodology

Data

I collected detailed point-by-point data from selected matches played at seven Grand Slam tournaments in 2006 and 2007.[12] Grand Slam tournaments are the most important and

[12] These are the seven Grand Slam tournaments going from the 2006 French Open (played in May-June 2006) to the 2007 US Open (played in August-September 2007).

prestigious tournaments on the professional tennis circuit. Each Grand Slam tournament has 128 entrants per gender, organized in a predetermined draw of 64 matches: the winner of a match advances to the next round, while the loser exits the tournament. The data were collected between July 2006 and August 2008 from the official web sites of the tournaments. One advantage of focusing on the Grand Slam tournaments is the uniformity of the available statistics, kept by IBM.

Detailed point-by-point data is available for a selected number of matches that were played on the main championship courts and were covered by IBM's Point Tracker technology.[13] The data was organized in two sets of files. The first set of files (one for each match) records for every point played in the match, who won the point, whether the first serve was in, the way the point ended (winner, unforced error, forced error, ace, double fault), and the score of the match. The second set of files (one for each point) tracks the movement of the ball throughout the rally in x-y-z coordinates, and allows to infer the number of strokes in each rally. I checked the data for consistency and only kept those points in which the number of strokes played is consistent with the information on the winner of the point and the way the point ends (for example, if the point ends in an unforced error by the server, the winner of the point must be the player returning serve and the number of strokes must be an odd number).[14]

For every point, all the strokes but the last one were classified as "put the ball in play". The last stroke was classified as either a winning shot or an unforced error, depending on how

[13] The Point Tracker technology was discontinued after the 2007 US Open.

[14] I was able to achieve a high level of consistency in the 2007 tournaments. For the 2006 tournaments, the level of consistency was lower, and a higher proportion of points had to be discarded. However, careful examination of the data revealed that many of the inconsistencies occurred in points where the number of strokes was off by one. I therefore prepared a second data set that includes also these potentially inconsistent observations, and verified that the main results are robust to use of this expanded data set.

the point ended. For points that ended in a forced error, the last stroke was dropped from the analysis, and the second to last stroke was classified as a winning shot.

I end up with a data set with information on more than 230,000 strokes in 52,000 points, distributed among 425 matches (227 men's matches and 198 women's matches). Finally, I also collected aggregate set-level data for *all* the matches played in the twelve Grand Slam tournaments going from the 2005 Australian Open to the 2007 US Open, and the ranking information of each player at the beginning of the tournament.

The Importance of Points

The key objective of the analysis is to construct a measure of the importance of each point and to study how performance varies with importance. Following Morris (1977), and Klaassen and Magnus (2001), I define the importance of a point as the probability that player 1 wins the match conditional on him or her winning the current point minus the probability that player 1 wins the match conditional on him or her losing the current point:

$$Importance_t = \text{Prob (player 1 wins match | player 1 wins point } t) - \text{Prob (player 1 wins match | player 1 loses point } t).$$

It is immediate to see that the importance of a point from the perspective of player 2 is exactly identical to the importance of a point from the perspective of player 1.

To calculate the importance of each point, I assume that in every match there is an associated fixed probability of each player winning a point, which depends on the gender, the playing surface, the two players' ability ratings and the identity of the server. Specifically, I use the data from the 5 tournaments not covered in the point-by-point analysis (the four tournaments played in 2005 and the 2006 Australian Open) to estimate the following regressions,[15] separately by gender:

$$\text{Proportion Points Won}_{ijsT} = \alpha_{sT} + \beta_{1s} Rating_{iT} + \beta_{2s} Rating_{jT} + u_{ijsT},$$

where the dependent variable is the proportion of points won by player *i* against player *j*, given serving status *s* and tournament *T*, and $Rating_{iT}$ and $Rating_{jT}$ are the *ability ratings* of players *i* and *j* respectively. The ability ratings are defined as $Rating = 8 - log_2(Rank)$. Klaassen and Magnus justify the use of this variable as a smoothed version of the expected round to be reached by a player of a given rank: for example, the number 1 ranked player in the world is expected to win all matches, and therefore reach round 8 (i.e., will win the tournament). This variable has three additional advantages: first, the distribution of this variable is less skewed than the distribution of rank, and it explains about twice the variance in the percentage of points won than the simple rank; second, it captures the fact that the difference in ability between the number 1 and the number 2 ranked players is probably greater than the difference between players ranked 101 and 102; finally, it takes on higher values for better players, which makes it easier to interpret it as a measure of ability. The estimated coefficients from the above regression are presented in Appendix Table 1, and conform to expectations: men are more likely to win points on serve, the server's advantage is larger on fast playing surfaces, and the ability gradient is larger for women than for men.

These regression coefficients are then used to fit the probabilities of winning a point in the detailed point-by-point data. These probabilities are then fed into a dynamic programming algorithm that takes into account the structure of a match in a Grand Slam tournament, and

[15] The proportion of points won is always bounded well away from 0 and 1, so that using a linear regression is appropriate in this case. Predicted probabilities are very similar if one uses a logistic regression instead.

calculates recursively, for every pairing of players, the probability of winning the match at every possible stage.[16] From this procedure it is then possible to calculate the importance of each point.

Table 2 presents summary statistics for the importance variable, separately by tournament and gender. The mean of the importance variable is 0.0243 for men, and 0.0282 for women. On average, points in the women's game are more important because matches are played in the "best-of-three" sets format, rather than "best-of-five." The distribution of the importance variable is heavily skewed to the right, indicating that most points played in a tennis match have relatively little potential to significantly affect the fate of a match.

The importance variable is able to identify effectively the points which any casual observer would think are indeed crucial for the final outcome of the match. This is shown in Table 3, which presents the average of the importance variable by set, status in the set, and status in the game. For example, points in the 5th set (average importance = 0.086) are on average about 5 times more important than points in the first set (average importance = 0.019). Importance also depends on whether the point is played at the early or late stages of the set, and on how evenly fought the sets are. The average of the importance variable when the score is 5-5 (in games) is 0.0406, about twice as large as when the score is 0-0 (0.0206), and almost 30 times as large as when the score is 5-0 (0.0015). There is also substantial variation within games: the average of the importance variable when the score is 40-0 is 0.0037, compared to 0.0534 when the score is 30-40.

[16] As an example, consider the probability of winning a game when the probability of winning a point is q and the score is "deuce" (40-40). This probability is equal to the probability of winning the next two points, plus the probability of winning only one of the next two points multiplied by the probability of winning the game when the score is "deuce" again. That is, $P_{40-40} = q^2 + 2q(1-q)P_{40-40} \Rightarrow P_{40-40} = q^2 / (1 - 2q(1-q))$. This calculated probability can then be used to calculate recursively $P_{40-30}, P_{30-30}, \ldots$

By construction, importance is larger the smaller the difference in ability between the two players. This makes intuitive sense: if the top ranked player in the world faces a breakpoint against the second ranked player, that point is much more likely to have an important effect on the final outcome of the match than if the top ranked player were facing a much lower-ranked player. This is because in an unequal contest, the top ranked player is expected to win a larger share of all subsequent points, and hence she will erase quickly the handicap of losing her service game.

The high nonlinearity of importance is also shown in Figures 1 and 2, which show the evolution of the importance variable over the course of the four finals played in the 2006 French Open and Wimbledon tournaments. Note how the importance variable evolves in a very nonlinear fashion: importance tends to rise towards the latter stages of each set, but only if the set is evenly fought. There are a number of clusters of high importance points even in the early sets and in the early stages of the late sets. Most of the spikes in importance are associated with break points. This is particularly true in the men's tournament at Wimbledon, where the fast playing surface means break points are relatively rare, and hence can change the direction of a match substantially.

Summing up, there is substantial variation in the importance measure both across matches and within matches, which should allow to detect variation in performance that depends on the degree of competitive pressure.

Econometric Methodology

I estimate the model parameters by maximum likelihood. Let T_j be the total number of strokes on point *j*, DU_j be a dummy variables indicating whether point *j* ends in an unforced error by either player, and DW_j be a dummy variables indicating whether point j ends in an unforced error by either player. Also, define u_{jt} as the probability of an unforced error on stroke t, point j, and w_{jt} as the probability of a winning shot on stroke t, point j. Then, the likelihood contribution of point j is:

$$L_j = \left(\prod_{t=1}^{T_j - 1} \left(1 - u_{jt} - w_{jt}\right) \right) \left(w_{jT_j}\right)^{DW_j} \left(u_{jT_j}\right)^{DU_j}$$

The probabilities of hitting winning shots and unforced errors are defined as follows:

$$u_{jt} = \frac{\exp\left(IU_{jt}\right)}{1 + \exp\left(IU_{jt}\right) + \exp\left(IW_{jt}\right)};$$
$$w_{jt} = \frac{\exp\left(IW_{jt}\right)}{1 + \exp\left(IU_{jt}\right) + \exp\left(IW_{jt}\right)},$$

with

$$IU_{jt} = \beta_2^u imp\_q2_{jt} + \beta_3^u imp\_q3_{jt} + \beta_4^u imp\_q4_{jt} + X'_{jt}\gamma^u;$$
$$IW_{jt} = \beta_2^w imp\_q2_{jt} + \beta_3^w imp\_q3_{jt} + \beta_4^w imp\_q4_{jt} + X'_{jt}\gamma^w.$$

IU_{jt} and IW_{jt} are linear index functions that include three quartiles of the importance variable; the ability rating of the player hitting the ball on stroke t and of his/her opponent; the point's serial number within the match, to capture fatigue effects; a dummy for whether the player hitting the ball at stroke t is the server; a dummy for whether the first serve on point j was in; the interaction of these last two dummies; and a series of dummies for each one of the first 10 strokes of the rally (with strokes beyond that being the omitted category); and tournament dummies to capture differences in type of surface, weather, and other fixed tournament characteristics. Some specifications will also include match fixed effects to capture differences in specific matchups between two given players, as well as any other unobservables that do not vary within matches (weather, temperature, time of day, relative cumulated fatigue of two players, etc.).

With this structure, it is clear that the model can be estimated using standard multinomial logit software. Standard errors in all specifications are clustered at the match level.

5. Results

Table 4 presents basic descriptive statistics on the percentage of points ending in unforced errors, on the number of strokes per point and on the typology of strokes by importance quartile. We see that for both men and women there is a moderate increase in the proportion of points ending in unforced errors between the first and fourth importance quartiles. The increase in the number of points ending in unforced errors could be the result of players adopting a more aggressive strategy and having a higher probability of hitting unforced errors on every stroke, but it could also occur if players are adopting a *less* aggressive strategy and reducing the number of winning shots.

The second panel in the table shows that the second hypothesis is more likely. The number of strokes per point increases sharply with importance quartile, for both men and women. The number of strokes per rally is between 15 and 21 percent higher in the fourth quartile of importance points relative to the first quarter. This suggests that players are adopting a less aggressive strategy as points become more important, playing long rallies rather than trying to win the point in few strokes.

The third panel in the table uses the stroke-level data and confirms this last point. There is a sharp and monotonic decrease in the percentage of *strokes* that are winning shots or unforced errors for both men and women. Note that while the ratio of unforced errors to winning shots is higher for women than for men, both men and women have a higher percentage of winning shots than unforced errors. Proposition 1 showed that in this case a decrease in the share of both winning shots and unforced errors implies a lower probability of winning the point. Hence, this first pass at the data seems to show that performance deteriorates on more important shots for both men and women. These results could of course be due to composition effects: maybe the more important points are disproportionately more likely to involve low-ranked players (who are more likely to commit unforced errors and less likely to hit winning shots), or are more likely to be played at the French Open, where the slow surface makes all rallies longer. To address these concerns, I proceed to the estimation of the full structural model described in the previous section.

Table 5 presents the results of the basic analysis. The model is estimated with and without match fixed effects. In the model without match fixed effects (columns 1 to 3), identification comes from variation in importance both between and within matches, while inclusion of fixed effects implies that the parameters of interest are identified solely off the variation in the importance variable within matches. The model is estimated separately for men (columns 1 and 4) and for women (columns 2 and 5). Columns 3 and 6 (the "difference" columns) are obtained by estimating the model on the joint sample of men and women, and interacting all the explanatory variables with a gender dummy, so that the point estimate is exactly equal to the difference between the two columns that precede it.

The top panel presents the coefficients on the importance variables in the index function for the probability of hitting winning shots, and the bottom panel shows the effect of importance on the probability of hitting unforced errors. For both men and women, there is clear evidence that the probability of hitting winners and unforced errors declines steeply with importance. For men, in the model without fixed effects, the odds of hitting a winner (relative to putting the ball in play) are about 28 log points lower in the top importance quartile than in the bottom importance quartile, and the odds of hitting an unforced error are 33 log points lower. The coefficients become even larger in the model with fixed effects. The pattern of coefficients for women is similar. The odds of hitting winners in the fourth importance quartile are between 19 and 44 log points lower than in the first importance quartile, and the odds of hitting unforced errors are between 14 and 29 log points lower. The third and sixth columns show the difference between the male and female coefficients. There is never a statistically significant difference between men and women in the probability of hitting winning shots. On the other hand, the decline in the women's probability of hitting unforced errors on important points is statistically smaller than the decline in the men's probability. For women, the change in the style of play between the top and bottom importance quartiles is smaller than for men, but the relative incidence of unforced errors increases. This can also be seen by looking at the difference between β_4^u and β_4^w (reported in the next row in the table), which tells us the effect of moving from the first to the fourth importance quartile on the ratio *u/w*. The statistic is close to zero for men, but positive for women (and statistically significant in the specification with match fixed effects). Thus, it would appear that women's performance deteriorates as we move from the first to the fourth importance quartile. However, as illustrated in Section 2, to assess precisely how performance changes, one must calculate the effect of changes in the probability of hitting unforced errors and winning shots on the probability of winning a point, for both genders.

This is implemented in Table 6, where I calculate the probability of winning a point against an "average" player. The "average" player is defined as one whose probability of hitting winning shots and unforced errors is equal to the sample average. I use the coefficients in Table 5 to calculate the probability of hitting winners and unforced errors by importance quartile, and I calculate the probability of winning a point as:

$$P(\text{A wins point} \mid \text{importance quartile } q) =$$
$$w_{q1} + p_{q1}\bar{u}_2 + p_{q1}\bar{p}_2 w_{q3} + \ldots + \prod_{t=1, t \text{ odd}}^{10} p_{qt} \prod_{t=1, t \text{ even}}^{10} \bar{p}_t \left(\frac{w_{q,11} + p_{q,11}\bar{u}_{12}}{1 - p_{q,11}\bar{p}_{12}} \right),$$

where w_{qt} and p_{qt} are the probabilities of hitting winning shots and returnable balls in importance quartile q and stroke t, and $\bar{u}_t$ and $\bar{p}_t$ are the probabilities of hitting unforced errors and returnable balls for the average player on stroke t. The probabilities are calculated separately by gender, serving status and first serve status. I then calculate the average probability of winning a point by serving status (averaged over first serve status, assuming that the probability of hitting the first serve in is independent of importance quartile and equal to the sample average), and report these probabilities (separately by gender) in Table 6.

In the model without fixed effects, the difference in the probability of winning a point between the first and fourth importance quartiles is relatively small: between 1.8 and 2.3 percentage points for servers, and less than 1 percentage point for receivers. The difference increases substantially in the model with fixed effects. The server's probability of winning the point falls by more than 4 percentage points for both men and women, and the receiver's probability of winning the point decreases by 1.2-2.3 percentage points. These differences are always statistically significant, and are quite substantial. For comparison purposes, Table 7 presents a matrix showing the proportion of points won by the server, by own rank and opponent's rank. For men, the proportion of points won by players ranked within the top 10 is between 4.6 and 6 percentage points higher than that of players ranked outside of the top 10. For women, the difference is between 4.6 and 7.1 percentage points. Therefore, changes in performance between importance quartiles are equivalent to somewhere between 50 and 100 percent of the difference in ability between top-10 and non top-10 players.

Figures 3 and 4 present an alternative way of looking at the relationship between importance and performance. The figures focus exclusively on variations in performance within *games*, and illustrate more transparently the relationship between performance and the importance variable. Specifically, I estimated a model similar to that in Table 5, replacing the importance variable with a series of dummy variables for all the possible combinations of points of server and receiver within a single game (0-0, 15-0, etc.). I then calculated the predicted probabilities of hitting winning shots and unforced errors for a representative match.

Figures 3a and 3b plot the odds of unforced errors relative to winning shots (i.e., the *u/w* ratio) at every combination of points within the game against the average value of the importance variable at that combination. The figure shows that for men there is only a weak correlation between the average importance of a point and the *u/w* ratio. For women, on the other hand, importance explains more than 40 percent of the variation in the *u/w* ratio. This confirms our previous findings that the incidence of unforced errors increases with importance for women more than it does for men.

Figures 4a and 4b plot the probability of winning the point against an "average" player implied by the estimated probabilities of unforced errors and winning shots. Here we see that for both men and women the fitted line has a strong negative slope, and importance can explain between 48 and 62 percent of the variation in the probability of winning the point. Both men and women exhibit a substantial deterioration in performance as points become more important.

Robustness Checks

I explore the robustness of the results to different sample definitions and specifications in Table 8. Panel A I use an extended data set that includes also about 14,000 points (60,000 strokes) from the 2006 tournaments that were dropped from the main sample because of an apparent inconsistency between the number of strokes in the rally, the way the point ended, and the winner of the point. Panel B uses an alternative calculation of the importance variable that is based only on the status of the match, and not on player abilities (hence, the probability of the number 1 ranked player to win the match at its very outset is always 0.5, regardless of his or her opponent is ranked 2 or 100). Finally, Panel C allows all the coefficients (save for those on the importance variables) to vary by tournament, recognizing the fact that ability ratings, fatigue effects and serve effects may all vary depending on the type of surface and on other atmospheric characteristics.Each panel presents first the coefficient on the fourth importance quartile in the winning shot and unforced error equations; then the difference between these two coefficients; and finally the implied difference in the probability of winning a point against an average player (as in Table 6).

The results in all three panels are almost identical to those obtained in the original sample: both men and women exhibit a decline in the number of winning shots and unforced errors as importance grows; for women, the decline in the share of unforced errors is smaller, meaning that the ratio of unforced errors to winning shots increases with importance; and the implied probability of winning the point against an average player declines substantially for both men and women, especially in the specifications with fixed effects. The only notable differences are in Panel B, in the specification without match fixed effects, where we find a substantially larger decline in the probability of winning shots and unforced errors for women.

In sum, it appears that our basic findings are quite robust to differences in sample selection, definition of the importance variable, and specification of the econometric model. In what follows, I explore the implications of the deterioration in performance on important points on the probability of winning a match.

6. Simulations

In this section, I analyze the effect of the deterioration in performance on important points on the probability of winning the match. Specifically, I ask what would happen to a player's probability of winning the match if he or she could keep on playing on more important points at the same level as on the less important points. I proceed as follows. I first simulate a match between two players of identical ability, whose probabilities of hitting winning shots and unforced errors are those implied by the coefficients from the fixed effects model in Table 5: on points with a value of the importance variable below the 75^{th} percentile, the probabilities of hitting winning shots and unforced errors are equal to the probabilities associated with the lowest importance quartile; on points with a value of the importance variable above the 75^{th} percentile, the probabilities are those associated with the highest importance quartile. The importance of each point is calculated assuming that players have identical abilities and play always according to the probabilities in the low importance quartile. These are the benchmark probabilities, shown in the first and third rows of the table. By construction each player's probability of winning the match is 0.5, because the two players have identical abilities.

I then run a second set of simulations (the "treatment" simulations), in which one of the two players does not experience any deterioration in performance on more important points (i.e., the probabilities of winning shots and unforced errors are the same on all points, regardless of importance). The results show that a player who is able to prevent the deterioration in performance on more important points could raise the probability of winning the match from

50% to 74-76% (for men), and to 81-83% (for women). Regressions similar to those in Appendix Table 1 indicate that a 1-point increase in ability rating is associated with roughly a 10 percentage point increase in the probability of winning the match. Therefore, a 30 percentage point increase in the probability of winning the match is equivalent to a three point increase in ability rating, or an eightfold change in the player's ranking. In other words, a player that could prevent the deterioration in performance on important points would dramatically increase his or her chances of advancing in the tournament, and reap the implied benefits in terms of prize money and ranking points.

7. Conclusion

In this paper I have used data from seven Grand Slam tournaments played between 2006 and 2007 to assess whether men and women respond differently to competitive pressure in a real-world setting with large monetary rewards. The analysis uses detailed stroke-by-stroke data and indicates that both men and women adopt a substantially less aggressive playing strategy on more important points. While there are significant gender differences in the relative incidence of unforced errors at crucial junctures of the match, these differences do not imply that women exhibit a more severe deterioration in performance on more important points. Our preferred metric for performance, the probability of winning the point against an "average" player, declines by roughly the same amount for both men and women. By way of a simple game-theoretic model, I argue that a switch to a safer playing style is likely to be the result of a decrease in the effectiveness of the aggressive strategy, and that players that can prevent the deterioration of performance on important points can substantially increase their chances of winning.

Summing up, the results support the notion that in certain situations high-powered incentives may have perverse effects because performance may actually decrease as the stakes become higher. On the other hand, I do not find much support for the claim that the underrepresentation of women at the upper echelons of management, science and academia can be attributed to differences in performance in competitive environments, and in particular to differences in the ways men and women respond to high pressure situations.

Of course, one should be careful in extrapolating from this study and draw more general lessons about how high-stakes rewards affect performance in the labor market. In fact, some specific features of tennis, such as different patterns of selection between elite male and female professional players, the fact that tasks involving motor skills (as opposed to tasks involving cognitive skills) may generate different responses to increases in competitive pressure, and the nature of the high pressure situations faced by tennis players (accurate decision-making and execution is needed in a matter of split seconds), should alert us to the dangers of extrapolation. Nevertheless, the fact that we find such robust evidence of underperformance in high-stakes situations, even among experienced professional drawn from the extreme right tail of the talent distribution, is without a doubt a novel finding that raises some interesting questions about the effect of high-powered incentive contracts, and should stimulate further investigation.

A1. Proofs and Extensions

Proof of Proposition 2. The proof proceeds in a number of steps. I first show that in any subgame perfect equilibrium in which both players play Soft at every stroke – call this the (Soft,Soft) equilibrium – a certain inequality must hold regarding the basic parameters u_0, u_1, w_0, and w_1. I then show that in any (Aggressive, Aggressive) equilibrium, a second inequality must hold. The third step proves that the two inequalities can never hold simultaneously. If the two conditions hold with equality, we are in the degenerate case in which player are always indifferent between playing Soft and Aggressive. Finally, it is shown that there can be no subgame perfect equilibrium in which one of the players plays Soft and the other plays Aggressive.

Lemma 1: If the pair of stationary strategies (Soft, Soft) is a subgame perfect equilibrium, it must be the case that:

$$\left(w_1 - w_0\right) + u_0\left(p_1 - p_0\right) + p_0\left(w_0 p_1 - w_1 p_0\right) < 0,$$

which can also be rewritten as:

$$\textbf{C1: } u_1 > \frac{u_0^2 + w_0\left(w_1 - w_0\right) + \left(1 - u_0 - w_0\right) w_1 u_0}{u_0 + \left(1 - u_0 - w_0\right) w_0}.$$

Proof: The pair of strategies (Soft, Soft) is a subgame perfect equilibrium if: a) $V^A(0,0) > V^A(1,0)$ and $V^B(0,0) > V^B(0,1)$; and b) $U^A(0,0) > U^A(1,0)$ and $U^B(0,0) > U^B(0,1)$. Part a) states that it is not optimal for one player to choose Aggressive at every stroke given that the other player chooses Soft at every stroke; part b) states that a one-time deviation from Soft is also not optimal. Since the two players are identical, it is enough to check the conditions for one of the two players. Now:

$$V^A(0,0) = w_0 + p_0 u_0 + p_0^2 V^A(0,0),$$
$$\Rightarrow V^A(0,0) = \frac{w_0 + p_0 u_0}{1 - p_0^2};$$

and:

$$V^A(1,0) = w_1 + p_1 u_0 + p_1 p_0 V^A(1,0),$$
$$= \frac{w_1 + p_1 u_0}{1 - p_1 p_0}.$$

Hence $V^A(0,0) > V^A(1,0)$ iff

$$\frac{w_0 + p_0 u_0}{1 - p_0^2} > \frac{w_1 + p_1 u_0}{1 - p_1 p_0}.$$

Rearranging terms, we have:

(1a) $$(w_1 - w_0) + u_0(p_1 - p_0) + p_0(w_0 p_1 - w_1 p_0) < 0.$$

Similarly, we have that

$$U^A(0,0) = w_0 + p_0 u_0 + p_0^2 \max\{V^A(0,0), V^A(1,0)\}$$
$$= w_0 + p_0 u_0 + p_0^2 V^A(0,0).$$
$$\Rightarrow U^A(0,0) = \frac{w_0 + p_0 u_0}{1 - p_0^2};$$

and:

$$U^A(1,0) = w_1 + p_1 u_0 + p_1 p_0 \max\{V^A(0,0), V^A(1,0)\}$$
$$= w_1 + p_1 u_0 + p_1 p_0 V^A(0,0)$$
$$= w_1 + p_1 u_0 + p_1 p_0 \frac{w_0 + p_0 u_0}{1 - p_0^2}.$$

Hence, $U^A(0,0) > U^A(1,0)$ iff

$$\frac{w_0 + p_0 u_0}{1 - p_0^2} > w_1 + p_1 u_0 + p_1 p_0 \frac{w_0 + p_0 u_0}{1 - p_0^2}.$$

Rearranging terms, we have:

(1b) $$(w_1 - w_0) + u_0(p_1 - p_0) + p_0(w_0 p_1 - w_1 p_0) < 0,$$

which is exactly the same as inequality (1a). After substituting $p_1 = 1 - u_1 - w_1$ and $p_0 = 1 - u_0 - w_0$, and some algebraic manipulations, this inequality can be rewritten as:

C1: $$u_1 > \frac{u_0^2 + w_0(w_1 - w_0) + (1 - u_0 - w_0) w_1 u_0}{u_0 + (1 - u_0 - w_0) w_0}. \blacksquare$$

Lemma 2: If the pair of stationary strategies (Aggressive, Aggressive) is a subgame perfect equilibrium, it must be the case that:

$$(w_1 - w_0) + u_1(p_1 - p_0) + p_1(w_0 p_1 - w_1 p_0) > 0,$$

which can also be rewritten as:

$$\textbf{C2: } (1-w_0)u_1^2 + \left[(1-w_1)(w_0-u_0)\right]u_1 - w_1\left[(w_1-w_0)+u_0(1-w_1)\right] < 0 .$$

Proof: The pair of strategies (Aggressive, Aggressive) is a subgame perfect equilibrium if: a) $V^A(1,1) > V^A(0,1)$ and $V^B(1,1) > V^B(1,0)$; and b) $U^A(1,1) > U^A(0,1)$ and $U^B(1,1) > U^B(1,0)$. Part a) states that it is not optimal for one player to choose Soft at every stroke given that the other player chooses Aggressive at every stroke; part b) states that a one-time deviation from Aggressive is also not optimal. Since the two players are identical, it is enough to check the conditions for one of the two players. Now:

$$V^A(1,1) = w_1 + p_1u_1 + p_1^2V^A(1,1),$$
$$\Rightarrow V^A(1,1) = \frac{w_1 + p_1u_1}{1-p_1^2};$$

and:

$$V^A(0,1) = w_0 + p_0u_1 + p_0p_1V^A(0,1),$$
$$= \frac{w_0 + p_0u_1}{1-p_0p_1}.$$

Hence $V^A(1,1) > V^A(0,1)$ iff

$$\frac{w_1 + p_1u_1}{1-p_1^2} > \frac{w_0 + p_0u_1}{1-p_0p_1}.$$

Rearranging terms, we have:

(2a) $$(w_1-w_0) + u_1(p_1-p_0) + p_1(w_0p_1 - w_1p_0) > 0 .$$

Similarly, we have that

$$U^A(1,1) = w_1 + p_1u_1 + p_1^2 \max\left\{V^A(0,1), V^A(1,1)\right\}$$
$$= w_1 + p_1u_1 + p_1^2V^A(1,1).$$
$$\Rightarrow U^A(1,1) = \frac{w_1 + p_1u_1}{1-p_1^2};$$

and:

$$U^A(0,1) = w_0 + p_0u_1 + p_0p_1 \max\left\{V^A(1,1), V^A(0,1)\right\}$$
$$= w_0 + p_0u_1 + p_0p_1V^A(1,1)$$
$$= w_0 + p_0u_1 + p_0p_1\frac{w_1 + p_1u_1}{1-p_1^2}.$$

Hence, $U^A(1,1) > U^A(0,1)$ iff

$$\frac{w_1 + p_1 u_1}{1 - p_1^2} > w_0 + p_0 u_1 + p_0 p_1 \frac{w_1 + p_1 u_1}{1 - p_1^2}.$$

Rearranging terms, we have:

(2b) $$(w_1 - w_0) + u_1(p_1 - p_0) + p_1(w_0 p_1 - w_1 p_0) > 0,$$

which is exactly the same as inequality (2a). After substituting $p_1 = 1 - u_1 - w_1$ and $p_0 = 1 - u_0 - w_0$, and some algebraic manipulations, this inequality can be rewritten as:

C2: $(1 - w_0)u_1^2 + \left[(1 - w_1)(w_0 - u_0)\right]u_1 - w_1\left[(w_1 - w_0) + u_0(1 - w_1)\right] < 0$.■

Lemma 3: Conditions **C1** and **C2** cannot hold simultaneously.

Proof: We note that the left hand side of condition C2 is a quadratic expression in u_1, and the coefficient on the quadratic term is positive. Hence, the left hand side is a U-shaped function in u_1, which takes on negative values when $u_1 \in (u_{1L}, u_{1H})$. At $u_1 = 0$ the expression is equal to $-w_1[(w_1 - w_0) + u_0(1 - w_1)]$, a negative number because of the assumption that $w_1 > w_0$. It follows that the smaller root u_{1L} is negative and we can ignore it, since u_1 is a probability and must always be in the interval [0,1]. Hence, condition **C2** is satisfied when u_1 is smaller than some value u_{1H}. Setting $u_1 = \frac{u_0^2 + w_0(w_1 - w_0) + (1 - u_0 - w_0)w_1 u_0}{u_0 + (1 - u_0 - w_0)w_0}$ (the right hand side of condition **C1**), makes the left hand side of **C2** equal to zero. Therefore **C2** is satisfied when $u_1 < \frac{u_0^2 + w_0(w_1 - w_0) + (1 - u_0 - w_0)w_1 u_0}{u_0 + (1 - u_0 - w_0)w_0}$, **C1** is satisfied when $u_1 > \frac{u_0^2 + w_0(w_1 - w_0) + (1 - u_0 - w_0)w_1 u_0}{u_0 + (1 - u_0 - w_0)w_0}$, and clearly the two conditions can not hold simultaneously, which is what we wished to prove.■

Lemma 4: There can be no subgame perfect equilibrium in which one player plays Aggressive at every stroke and the other player plays Soft at every stroke.

Proof: Without loss of generality, I will prove that the pair of strategies (Aggressive, Soft) cannot be a subgame perfect equilibrium. Assume by contradiction that it is. Then $V^A(1,0) > V^A(0,0)$ and $V^B(1,0) > V^B(1,1)$. This implies that

$$\frac{w_1 + p_1 u_0}{1 - p_1 p_0} > \frac{w_0 + p_0 u_0}{1 - p_0^2},$$

and

$$\frac{w_0 + p_0 u_1}{1 - p_0 p_1} > \frac{w_1 + p_1 u_1}{1 - p_1^2}.$$

Rearranging, we have that (Aggressive, Soft) is a subgame perfect equilibrium if $(w_1 - w_0) + u_0(p_1 - p_0) + p_0(w_1 p_0 - w_0 p_1) > 0$. Note that this is exactly the opposite inequality of that needed for (Soft, Soft) to be an equilibrium. But Lemma 3 implies that when $(w_1 - w_0) + u_0(p_1 - p_0) + p_0(w_1 p_0 - w_0 p_1) > 0$, (Aggressive, Aggressive) is an equilibrium, or $V^B(1,1) > V^B(1,0)$. This is in contradiction with the stated assumption that (Aggressive, Soft) is an equilibrium and therefore it is not desirable for player B to deviate (i.e., $V^B(1,0) > V^B(1,1)$). Hence we have a contradiction, and (Aggressive, Soft) cannot be a subgame perfect equilibrium.■

Lemma 3 proves that the equilibrium can be either (Soft,Soft) or (Aggressive, Aggressive), but there can be no combination of parameters such that both pairs of stationary strategies are subgame perfect equilibria. The two inequalities in **C1** and **C2** are essentially identical. If they hold with equality, it means than both players are always indifferent between playing Soft and Aggressive, and therefore we are in the degenerate case where any combination of mixed strategies by the two players is a subgame perfect equilibrium. This completes the proof. *QED.*

Extension: heterogeneous abilities. If players have heterogeneous abilities, Propositions 2 can be modified as follows:

Proposition 2': In the game described in the text, when players have heterogeneous abilities, only one of five outcomes is possible: a) a unique pure-strategy subgame perfect equilibrium in which both players play "soft"; b) a unique pure-strategy subgame perfect equilibrium in which both players play "aggressive"; c) a unique pure-strategy subgame perfect equilibrium in which player A plays "aggressive" and player B plays "soft"; d) a unique pure-strategy subgame perfect equilibrium in which player A plays "soft" and player B plays "aggressive"; e) a degenerate situation in which one of the two players is indifferent between playing "soft" and "aggressive," and therefore an infinite number of mixed-strategy subgame perfect equilibria may arise.

Extension: Non-stationarity. I modify the set-up of the game described in the text by allowing a fourth type of shot, a forcing shot. Now, players A and B can choose to play Soft or Aggressive when they face a hittable shot, but are forced to play a defensive shot when they face a forcing shot. The probabilities that player j (j = A, B) hits a winner, an unforced error and a forcing shot are, respectively, w_{0j}, u_{0j}, f_{0j} when playing Soft, and w_{1j}, u_{1j}, f_{1j} when playing Aggressive. The probability of making an error when facing a forcing shot is e_j. We modify Assumption 1:

Assumption 1'. a) $w_{1j} > w_{0j}$, for j = A, B **;** b) $u_{1j} > u_{0j}$, for j = A, B **;** c) $f_{1j} > f_{0j}$, for j = A, B; d) $w_{1j}, w_{0j}, u_{1j}, u_{0j}, f_{1j}$ and f_{0j} are constant over the course of the rally.

Then, if players have identical abilities, Proposition 2 still holds. The (Soft, Soft) equilibrium will arise if the following condition holds:

$$\textbf{C1'}: \frac{(w_0 + f_0 e)[1-(1-e)f_0] + (1 - w_0 - u_0 - f_0)u_0}{[1-(1-e)f_0]^2 - (1 - w_0 - u_0 - f_0)^2} >$$

$$\frac{(w_1 + f_1 e)[1-(1-e)f_0] + (1 - w_1 - u_1 - f_1)u_0}{[1-(1-e)f_0][1-(1-e)f]_1 - (1 - w_0 - u_0 - f_0)(1 - w_1 - u_1 - f_1)}.$$

References

Ariely, Dan; Gneezy, Uri; Loewenstein, George and Mazar, Nina. "Large Stakes and Big Mistakes." Federal Reserve Bank of Boston Working Paper 05-11, July 2005.

Babcock, Linda and Laschever, Sara. *Women Don't Ask: Negotiation and the Gender Divide.* Princeton and Oxford: Princeton University Press, 2003.

Baumeister, Roy F. "Choking under Pressure: Self-consciousness and Paradoxical Effects of Incentives on Skillful Performance." *Journal of Personality and Social Psychology,* 46(3), March 1984, 610-620.

Baumeister, Roy F.; Hamilton, James C. and Tice, Dianne M. "Public versus Private Expectancy of Success: Confidence Booster or Performance Pressure?" *Journal of Personality and Social Psychology*, 48(6), June 1985, pp. 1447-57.

Beilock, Sian L. and Carr, Thomas H. "When high-powered people fail: Working memory and 'choking under pressure' in math." *Psychological Science*, 16(2), February 2005, pp. 101-05.

Bertrand, Marianne and Hallock, Kevin F. "The Gender Gap in Top Corporate Jobs." *Industrial and Labor Relations Review*, 55(1), October 2001, pp. 3-21.

Booth, Alison L. and Nolen, Patrick J. "Choosing to Compete: How Different Are Girls and Boys?" IZA Discussion Paper No. 4027, February 2009.

Butler, Jennifer L. and Baumeister, Roy F. "The Trouble with Friendly Faces: Skilled Performance with a Supportive Audience." *Journal of Personality and Social Psychology*, 75(5), November 1998, pp. 1213-30.

Croson, Rachel and Gneezy, Uri. "Gender Differences in Preferences." Mimeo., University of Pennsylvania, September 2004.

Deaner, Robert O. "More males run fast: A stable sex difference in competitiveness in U.S. distance runners." *Evolution and Human Behavior*, 27(1), January 2006a, pp. 63-84.

Deaner, Robert O. "More males run relatively fast in U.S. road races: Further evidence of a sex difference in competitiveness." *Evolutionary Psychology*, vol. 4, 2006b, pp. 303-14.

Dohmen, Thomas J. "Do Professionals Choke Under Pressure?" *Journal of Economic Behavior and Organization*, 2006, doi:10.1016/j.jebo.2005.12.004.

Dohmen, Thomas J. and Falk, Armin. "Performance, Pay and Multi-dimensional Sorting: Productivity, Preferences and Gender." IZA Discussion Paper No. 2001, March 2006.

Gneezy, Uri; Niederle, Muriel, and Rustichini, Aldo. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*, 118(3), August 2003, pp. 1049-74.

Gneezy, Uri and Rustichini, Aldo. "Gender and Competition at a Young Age". *American Economic Review,* 94(2), May 2004, pp. 377-81.

Goldin, Claudia and Rouse, Cecilia. "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians." *American Economic Review*, 90(4), September 2000, pp. 715-741.

Klaassen, Franc J.G.M., and Magnus, Jan R. "Are Points in Tennis Independent and Identically Distributed? Evidence from a Dynamic Binary Panel Data Model." *Journal of the American Statistical Association*, 96(454), June 2001, pp. 500-509.

Kleine, D.; Sampedro, R. and Lopes Melo, S. "Anxiety and performance in runners. - Effects of stress on anxiety and performance in ergometric tests and competitions." *Anxiety Research*, 1, 1988, pp. 235-246.

Lavy, Victor. "Do Gender Stereotypes Reduce Girls' Human Capital Outcomes? Evidence from a Natural Experiment." *Journal of Public Economics*, 92(10-11), October 2008a, pp. 2083-2105.

Lavy, Victor. "Gender Differences in Competitiveness in a Real Workplace: Evidence from Performance-Based Pay Tournaments among Teachers." NBER Working Paper 14338, September 2008b.

Manning, Alan and Saidi, Farzad. "Understanding the Gender Pay Gap. What's Competition Got to Do with It?" CEP Discussion Paper, 2008.

Massachusetts Institute of Technology. *A Study on the Status of Women Faculty in Science at MIT*, 1999. Downloadable from *http://web.mit.edu/fnl/women/women.pdf.*

Morris,C. "The Most Important Points in Tennis," in *Optimal Strategies in Sport*, eds. S.P.Ladany and R.E.Machol, 1977. Amsterdam: North-Holland Publishing Company, 131-140.

Niederle, Muriel, and Vesterlund, Lise. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics*, August 2007, forthcoming.

Paserman, M. Daniele. "Gender Differences in Performance in Competitive Environments: Evidence from Professional Tennis Players." CEPR Discussion Paper 6335, June 2007.

Summers, Lawrence H. *Remarks at NBER Conference on Diversifying the Science and Engineering Workforce*, January 2005. Transcript taken from *http://www.president.harvard.edu/speeches/2005/nber.html.*

Sunde, Uwe. "Potential, Prizes and Performance: Testing Tournament Theory with Professional Tennis Data." IZA Discussion Paper No. 947, December 2003.

Weinberg, Richard S.; Richardson, P.A. and Jackson, Allen. "Effect of situation criticality on tennis performance of males and females." *International Journal of Sport Psychology*, 12(4), October-December1981, pp. 253-259.

Table 1: Equilibrium Strategies and Outcomes for Selected Parameter Values

	Intrinsic probabilities when playing Soft (1)	Intrinsic probabilities when playing Aggressive	Equilibrium strategies (3)	Equilibrium probability of unforced error (4)	Equilibrium number of strokes (5)	P(A wins point) [serving, receiving] (5)
1	$w_0 = 0.13$ $u_0 = 0.08$	$w_1 = 0.15$ $u_1 = 0.09$	(Aggressive, Aggressive)	$\frac{u_1}{u_1 + w_1} = 0.375$	$\frac{1}{u_1 + w_1} = 4.16$	0.517, 0.483
2	$w_0 = 0.13$ $u_0 = 0.08$	$w_1 = 0.15$ $u_1 = 0.11$	(Soft, Soft)	$\frac{u_0}{u_0 + w_0} = 0.381$	$\frac{1}{u_0 + w_0} = 4.76$	0.514, 0.486
3	$w_{0A} = 0.13$ $u_{0A} = 0.08$ $w_{0B} = 0.13$ $u_{0B} = 0.08$	$w_{1A} = 0.15$ $u_{1A} = 0.09$ $w_{1B} = 0.15$ $u_{1B} = 0.11$	(Aggressive, Soft)	$\frac{u_{1A} + p_{1A}u_{1B}}{1 - p_{1A}p_{1B}} = 0.377$	$\frac{1 + p_{1A}}{1 - p_{1A}p_{1B}} = 4.40$	0.528, 0.497

Note: The table presents the equilibrium strategies and outcomes for a single point according to the model presented in Section 2.

Table 2: The Importance of Points – Summary Statistics

	Men	**Women**	**ALL**
Mean	0.0243	0.0282	**0.0256**
Standard deviation	0.0351	0.0386	**0.0364**
25th percentile	0.0027	0.0030	**0.0028**
50th percentile	0.0125	0.0136	**0.0129**
75th percentile	0.0324	0.0394	**0.0347**
Number of points	43,620	22,590	**66,210**
Number of points with consistent information on number of strokes and way points end	34,438	18,033	**52,471**
Number of matches	227	198	**425**

Note: The importance of a point is defined as the probability that player 1 wins the entire match conditional on him/her winning the current point, minus the probability that player 1 wins the entire match conditional on him/her winning the current point. See text for details.

Table 3: The Importance of Points, by Match, Set and Game Status

Mean importance by match status				Mean importance by set status				Mean importance by game status			
Score in sets	All	Men	Women	Score in games	All	Score in games	All	Score in points	All	Men	Women
2-0	0.0107	0.0107	.	5-0	0.0015	3-1	0.0217	40-0	0.0037	0.0027	0.0063
0-0	0.0194	0.0183	0.0208	0-5	0.0020	2-1	0.0233	40-15	0.0096	0.0080	0.0134
1-0	0.0215	0.0180	0.0257	1-5	0.0041	1-2	0.0240	30-0	0.0098	0.0081	0.0140
2-1	0.0379	0.0379	.	0-4	0.0061	1-1	0.0242	15-0	0.0162	0.0142	0.0205
1-1	0.0497	0.0357	0.0720	5-1	0.0077	4-2	0.0251	0-40	0.0189	0.0238	0.0133
2-2	0.0864	0.0864	.	2-5	0.0078	3-2	0.0274	0-0	0.0201	0.0188	0.0227
				5-2	0.0088	2-2	0.0275	30-15	0.0203	0.0182	0.0245
				4-0	0.0097	5-3	0.0279	40-30	0.0247	0.0208	0.0322
				1-4	0.0103	2-3	0.0281	15-15	0.0257	0.0244	0.0281
				0-3	0.0126	4-3	0.0328	0-15	0.0267	0.0269	0.0264
				4-1	0.0131	3-3	0.0337	0-30	0.0282	0.0305	0.0250
				3-0	0.0135	3-4	0.0343	15-40	0.0326	0.0369	0.0269
				1-3	0.0141	5-5	0.0406	15-30	0.0347	0.0354	0.0335
				0-2	0.0155	5-4	0.0414	deuce	0.0361	0.0343	0.0392
				2-4	0.0156	4-4	0.0430	30-40	0.0534	0.0563	0.0492
				3-5	0.0157	4-5	0.0453	TB	0.0744	0.0726	0.0834
				2-0	0.0199	5-6	0.0459				
				0-0	0.0206	6-5	0.0506				
				0-1	0.0206	TB	0.0744				
				1-0	0.0209	6-6	0.1331				

Note: The first number in the "score in games" column represents the number of games won by the server, the second number is the number of games won by the receiver.

Table 4: Typology of points and strokes by importance quartile

A: Percentage of points ending in unforced error

	Men	Women
Importance quartile 1	29.27	37.70
Importance quartile 2	30.34	38.09
Importance quartile 3	30.00	37.68
Importance quartile 4	30.66	38.35
Number of points	34,393	17,966

B: Number of strokes per point

	Men	Women
Importance quartile 1	3.97 (3.54)	4.21 (3.43)
Importance quartile 2	4.29 (3.89)	4.43 (3.71)
Importance quartile 3	4.48 (4.11)	4.42 (3.72)
Importance quartile 4	4.81 (4.48)	4.80 (4.04)
Number of points	34,393	17,966

C: Percentage of strokes that are winning shots/ unforced errors

	Men		Women	
	Winning shot	Unforced error	Winning shot	Unforced error
Importance quartile 1	17.81	7.37	14.80	8.96
Importance quartile 2	16.25	7.07	13.99	8.61
Importance quartile 3	15.60	6.71	14.09	8.52
Importance quartile 4	14.42	6.38	12.86	8.00
Number of strokes	149,917		80,375	

Note: Calculations based on all points for which there was no inconsistency between the number of strokes, the winner of the point, and the type of shot with which the point ended.

Table 5: The Effect of Importance on the Probability of Hitting Winning Shots and Unforced Errors

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
Winning shots						
Importance quartile 2 (β_2^w)	-0.097***	-0.059*	0.038	-0.185***	-0.163***	0.022
	(.030)	(.035)	(.046)	(.030)	(.039)	(.049)
Importance quartile 3 (β_3^w)	-0.147***	-0.049	0.098*	-0.301***	-0.221***	0.080
	(.035)	(.045)	(.057)	(.037)	(.048)	(.061)
Importance quartile 4 (β_4^w)	-0.279***	-0.192***	0.087	-0.454***	-0.436***	0.018
	(.039)	(.048)	(.062)	(.041)	(.058)	(.071)
Unforced Errors						
Importance quartile 2 (β_2^u)	-0.104***	-0.069	0.035	-0.164***	-0.082*	0.083
	(.035)	(.046)	(0.058)	(.040)	(.049)	(.063)
Importance quartile 3(β_3^u)	-0.187***	-0.065	0.122**	-0.304***	-0.128**	0.176**
	(.038)	(.049)	(0.061)	(.042)	(.055)	(.069)
Importance quartile 4(β_4^u)	-0.325***	-0.138***	0.187***	-0.459***	-0.288***	0.171**
	(.041)	(.049)	(0.063)	(.047)	(.060)	(.076)
Test of: $\beta_4^u - \beta_4^w = 0$						
	-0.046	0.055	0.100	-0.005	0.148**	0.153
	(.047)	(.063)	(.078)	(.060)	(.074)	(0.095)
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of points	34,438	18,033	52,471	34,438	18,033	52,471
Number of strokes	149,917	80,375	230,292	149,917	80,375	230,292

Note: Entries in the table are the coefficients on the importance quartile dummies in the structural model described in the text. Additional control variables: own rating, opponent's rating, serial number of the point within the match, server dummy, first serve dummy, server×first serve, flexible function of server and stroke within rally (see text), tournament dummies. Robust standard errors (adjusted for clustering at the match level) in parentheses.

Table 6: Implied Probability of Winning Point against "Average" Player, by Importance Quartile

	Men		Women	
	Server	Receiver	Server	Receiver
A: No Fixed effects				
Importance quartile 1	0.657	0.365	0.559	0.435
	(.019)	(.008)	(.023)	(.013)
Importance quartile 4	0.634	0.359	0.542	0.426
	(.019)	(.008)	(.021)	(.011)
Difference	-0.023	-0.006	-0.018	-0.009
	(.015)	(0.006)	(.016)	(.009)
B: Fixed Effects				
Importance quartile 1	0.656	0.365	0.591	0.456
	(.033)	(.013)	(.038)	(.021)
Importance quartile 4	0.615	0.353	0.547	0.434
	(.035)	(.013)	(.038)	(.020)
Difference	-0.041***	-0.012**	-0.043**	-0.023**
	(.016)	(.006)	(.019)	(.011)

Note: The probabilities of winning a point are calculated by using the implied probabilities (from Table 5) of winning shots and unforced errors for a representative player, assuming that the opponent's probability of hitting wining shots and unforced errors is equal to the sample average. See text for details.

Table 7: Probability of Server Winning Point, by Own Rank and Opponent Rank

	Men		Women	
	Opponent rank <=10	Opponent rank >10	Opponent rank <=10	Opponent rank >10
Own rank <=10	0.642	0.696	0.569	0.617
Own rank>10	0.598	0.636	0.498	0.571
Difference	**0.046**	**0.060**	**0.071**	**0.046**

Note: Probabilities calculated on the full sample of matches in the analysis sample.

Table 8: Robustness Checks

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
A:Extended data set						
Winning shots, importance quartile 4 (β_4^w)	-0.270*** (.037)	-0.176*** (.046)	0.094 (.059)	-0.443*** (.037)	-0.421*** (.050)	0.022 (.062)
Unforced errors, importance quartile 4(β_4^u)	-0.321*** (.039)	-0.096** (.048)	0.225*** (.062)	-0.463*** (.041)	-0.244*** (.055)	0.219*** (.068)
Test of: $\beta_4^u - \beta_4^w = 0$	-0.051 (.041)	0.080 (.060)	0.131* (.073)	-0.020 (.053)	0.176*** (.064)	0.197** (.083)
Implied difference in probability of winning point between importance quartiles 1 and 4						
Server	-0.022 (.014)	-0.020 (.015)	-	-0.039*** (.014)	-0.046*** (.016)	-
Receiver	-0.007 (.006)	-0.011 (.009)	-	-0.014** (.006)	-0.025*** (.010)	-
B: Importance not a function of player ability						
Winning shots, importance quartile 4 (β_4^w)	-0.346*** (.028)	-0.428*** (.039)	-0.082* (.047)	-0.352*** (.027)	-0.438*** (.038)	-0.086* (.047)
Unforced errors, importance quartile 4(β_4^u)	-0.341*** (.034)	-0.235*** (.040)	0.106** (.052)	-0.337*** (.034)	-0.240*** (.043)	0.097* (.055)
Test of: $\beta_4^u - \beta_4^w = 0$	0.006 (.041)	0.193*** (.051)	0.188*** (.065)	0.015 (.041)	0.198*** (.054)	0.183*** (.068)
Implied difference in probability of winning point between importance quartiles 1 and 4						
Server	-0.034*** (.011)	-0.047*** (.013)	-	-0.034*** (.011)	-0.049*** (.013)	
Receiver	-0.011** (.004)	-0.024*** (.007)	-	-0.011** (.004)	-0.026*** (.007)	

Note: Robust standard errors (adjusted for clustering at the match level) in parentheses.

Table 8: Robustness Checks (continued)

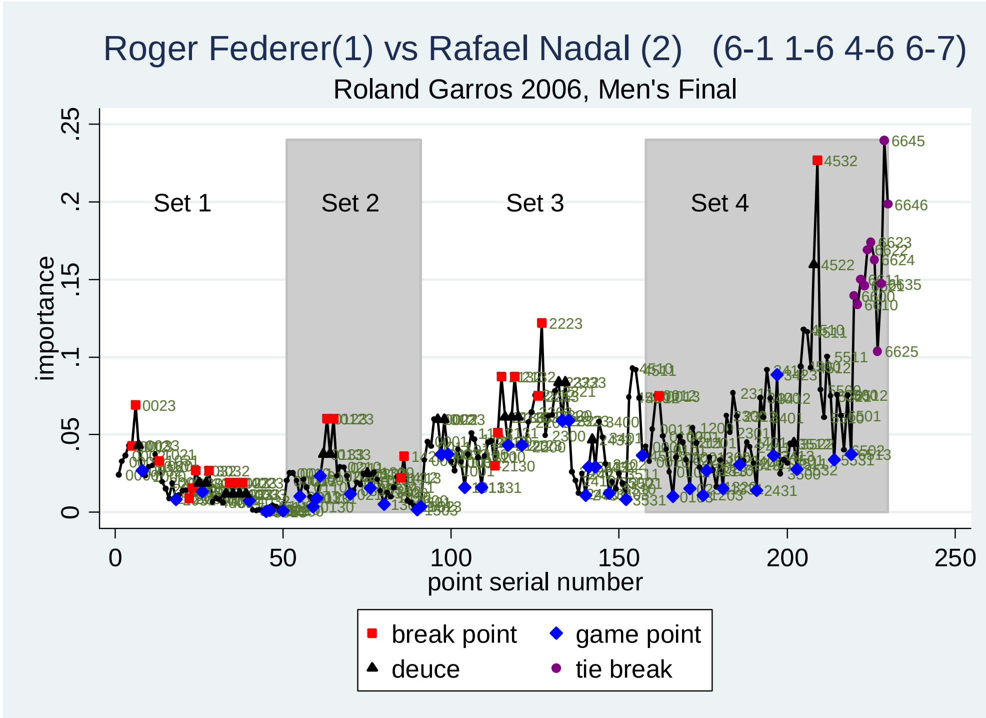
C: All coefficients interacted with tournament dummies						
Winning shots, importance quartile 4 (β_4^w)	-0.281*** (.039)	-0.191*** (.048)	0.091 (.061)	-0.453*** (.040)	-0.448*** (.056)	0.005 (.069)
Unforced errors, importance quartile 4(β_4^u)	-0.321*** (.041)	-0.142*** (.048)	0.179*** (.063)	-0.462*** (.047)	-0.306*** (.061)	0.156** (.077)
Test of: $\beta_4^u - \beta_4^w = 0$	-0.039 (.046)	0.049 (.063)	0.088 (.079)	-0.010 (.059)	0.142* (.074)	0.152 (.094)
Implied difference in probability of winning point between importance quartiles 1 and 4						
Server	-0.024 (…)	-0.017 (…)	-	-0.042 (…)	-0.048 (…)	-
Receiver	-0.007 (…)	-0.009 (…)	-	-0.017 (…)	-0.028 (…)	-

Note: Robust standard errors (adjusted for clustering at the match level) in parentheses.

Table 9: The Effect of Raising One's Game on Important Points on the Probability of Winning the Match

	Low Importance points Probability that player A wins point when:		High importance points Probability that player A wins point when:		Probability that player A wins match
	Serving	Receiving	Serving	Receiving	
A: Men					
Benchmark (Both players soft on low importance points, aggressive on high importance points)	0.6496	0.3504	0.6201	0.3799	**Best of 3 sets: 0.5** **Best of 5 sets: 0.5**
Treatment (Player A plays aggressive on important points)	0.6496	0.3504	0.6615	0.3947	**Best of 3 sets: 0.735** **Best of 5 sets: 0.763**
B: Women					
Benchmark (Both players soft on low importance points, aggressive on high importance points)	0.5728	0.4271	0.5498	0.4502	**Best of 3 sets: 0.5** **Best of 5 sets: 0.5**
Treatment (Player A plays aggressive on important points)	0.5728	0.4271	0.5941	0.4744	**Best of 3 sets: 0.807** **Best of 5 sets: 0.826**

Note: The rows denoted "benchmark" present the probabilities of winning a point in a match between two players of equal ability (benchmark case), when the two players' probabilities of hitting winning shots and unforced errors are equal to those implied by the coefficients in Table 5: quartile 1 for low importance points, quartile 4 for high importance points. The threshold is the 75th percentile of the importance variable in the sample. The importance of each point is calculated assuming that players have identical abilities and play always according to the probabilities in the low importance quartile. The rows denoted "treatment" show how the predicted probabilities change when player A's probability of hitting winning shots and unforced errors does not change on the more important points. All estimated probabilities are based on 5,000 simulated matches.



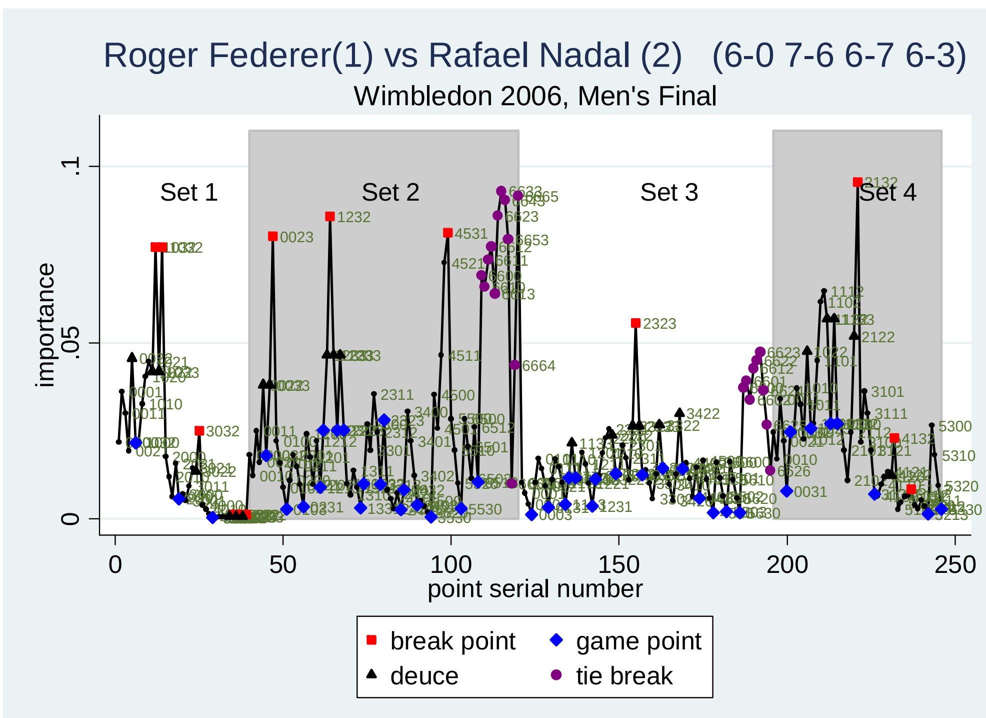
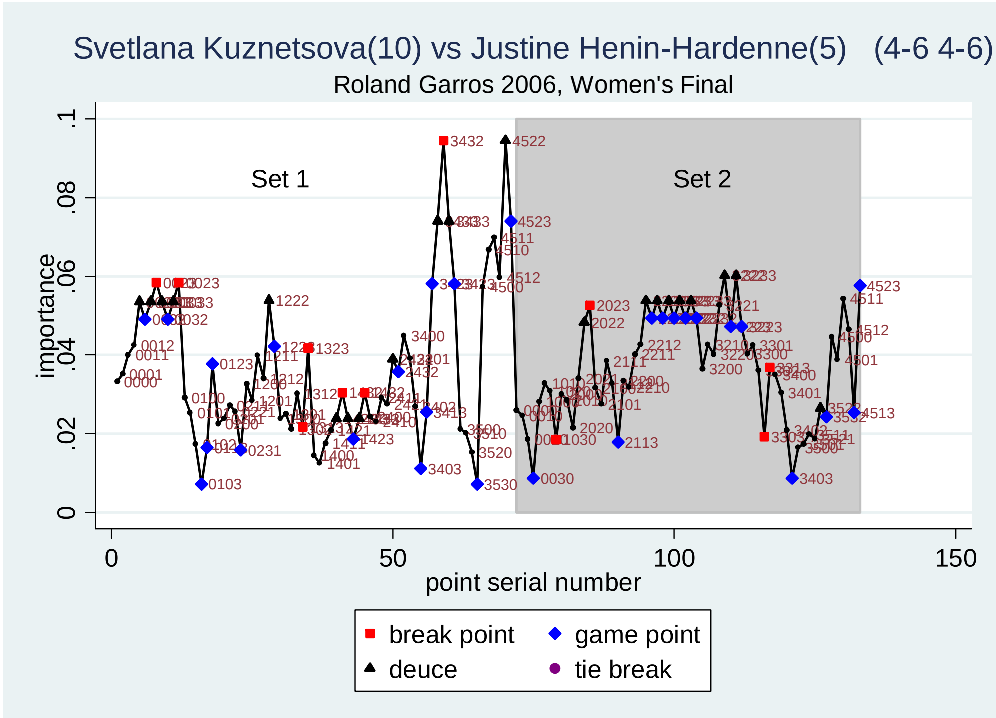


Figure 1: The evolution of importance over the course of selected matches – men

Note: The 4-digit string next to each label denotes the score within the set: games won by player 1, games won by player 2, points won by player 1 and points won by player 2.



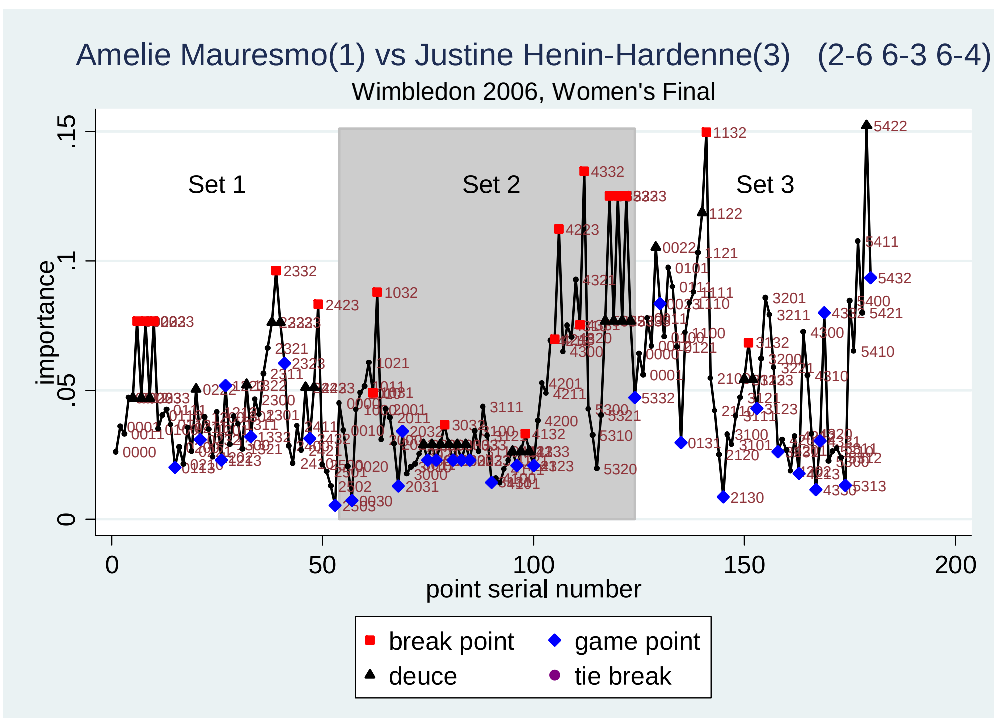


Figure 2: The evolution of importance over the course of selected matches – women

Note: The 4-digit string next to each label denotes the score within the set: games won by player 1, games won by player 2, points won by player 1 and points won by player 2.

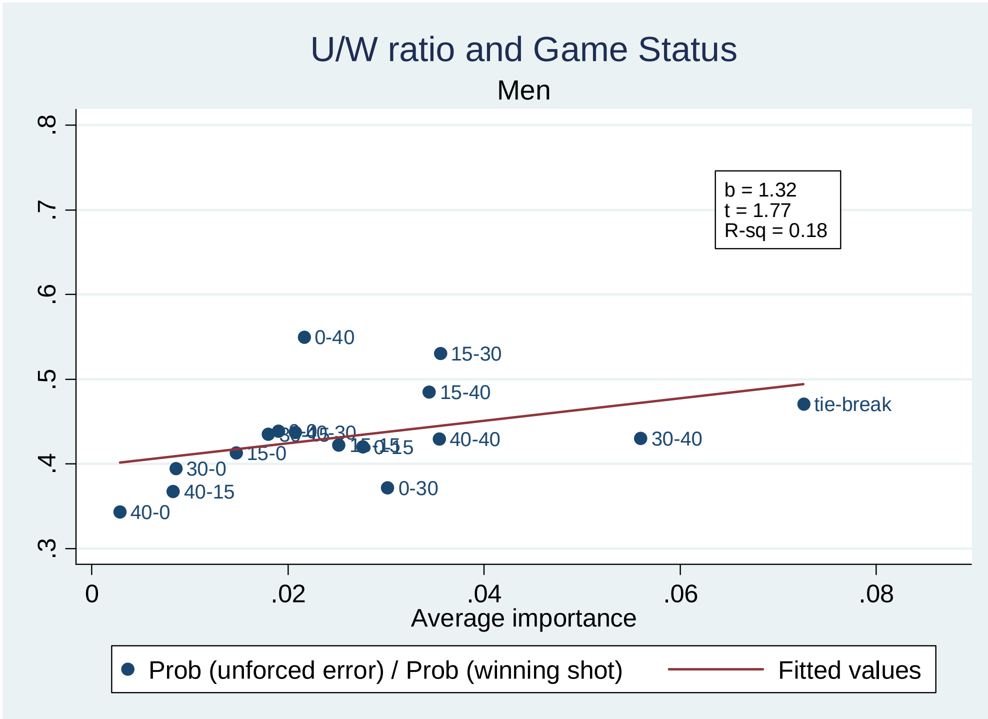


Figure 3a: Unforced Errors/Winning Shot Ratio and Game Status – Men

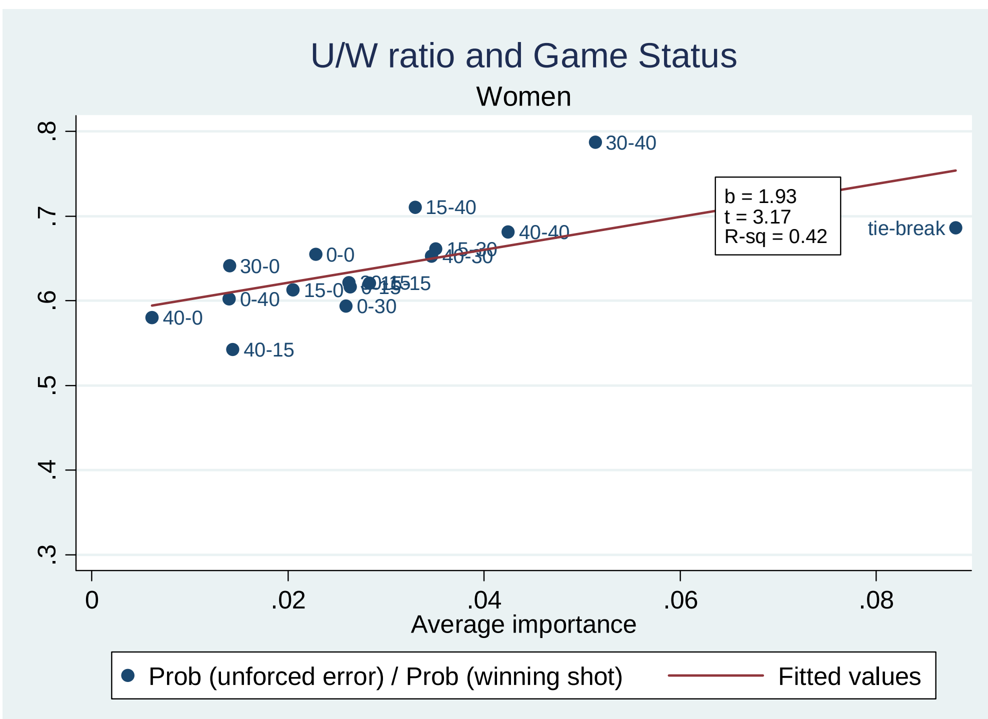


Figure 3b: Unforced Errors/Winning Shot Ratio and Game Status – Women

Note: On the vertical axis is the predicted ratio of unforced errors to winning shots (the u/w ratio) for a representative point; on the horizontal axis is the average of the importance value for the combination of points won by the server and by the receiver within a game. The fitted line is obtained from the simple regression of the predicted u/w ratio on average importance.

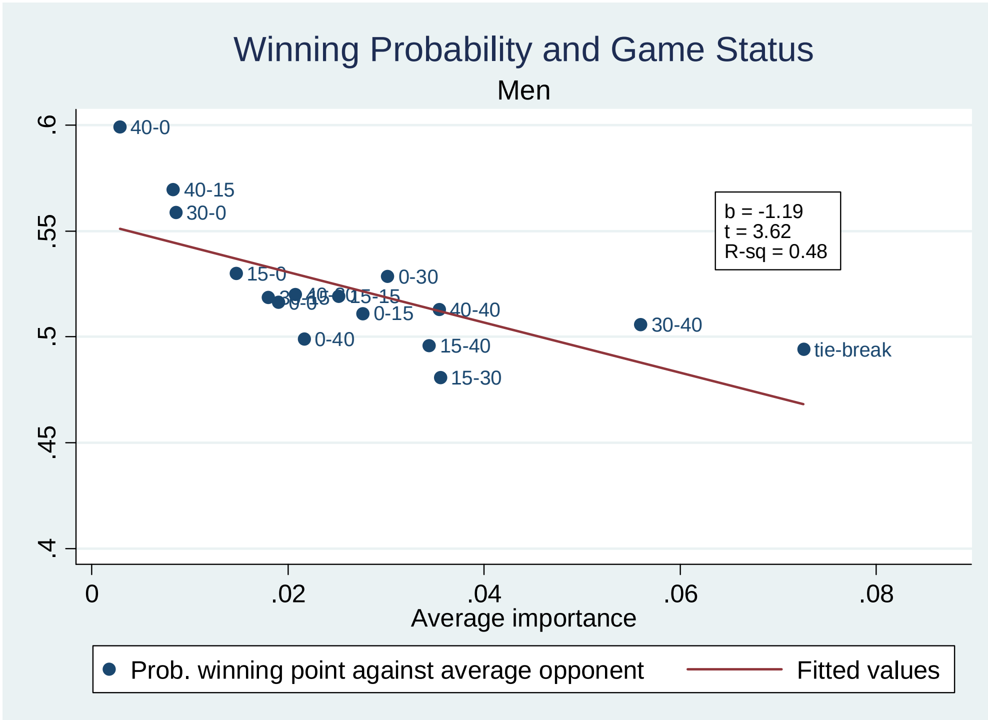


Figure 4a: Probability of winning the point and game status – Men

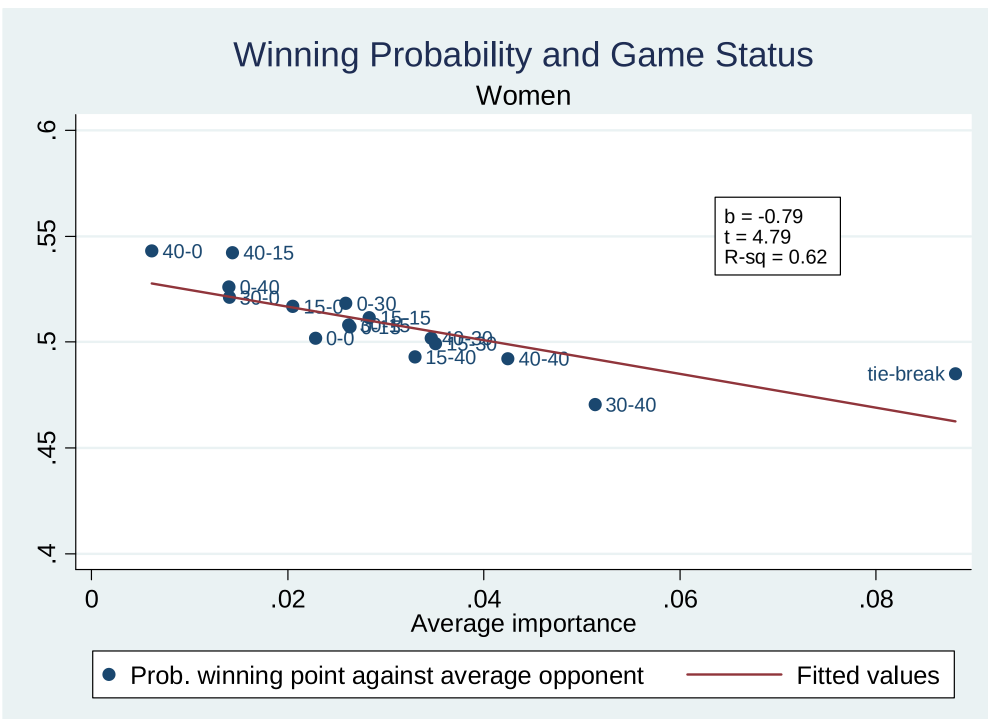


Figure 4a: Probability of winning the point and game status – Women

Note: On the vertical axis is the predicted probability of winning the point against an "average" player, implied by the predicted probabilities of hitting winning shots and unforced errors; on the horizontal axis is the average of the importance value for the combination of points won by the server and by the receiver within a game. The fitted line is obtained from the simple regression of the predicted probability on average importance.

Appendix Table 1: Estimating the Probability of Winning a Point

Dependent variable: probability of server winning the point

	Men	Women
Server's ability rating	0.0156 (0.0012)	0.0208 (0.0014)
Receiver's ability rating	-0.0140 (0.0012)	-0.0182 (0.0014)
Roland Garros	-0.0249 (0.0058)	-0.0173 (0.0069)
Wimbledon	0.0367 (0.0058)	0.0252 (0.0069)
US Open	0.0141 (0.0066)	0.0079 (0.0080)
Constant	0.6149 (0.0058)	0.5447 (0.0069)
N	1148	1174
R^2	0.259	0.265

Note: The table presents the coefficients from a regression of the proportion of points won by the server on the server's and the receiver's ability rating, and tournament dummies. Standard errors in parentheses. The regressions were estimated using data on the 2005 and 2006 Australian Opens, Roland Garros 2005, Wimbledon 2005, and the 2005 US Open.