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

M. Daniele Paserman
Hebrew University, CEPR, and IZA
dpaserma@shum.huji.ac.il

January 2007

Abstract

This paper uses data from all tennis Grand Slam tournaments played in 2005 and 2006 to assess whether men and women respond differently to competitive pressure in a setting with large monetary rewards. In particular, it asks whether the quality of the game deteriorates as the stakes become higher. The paper conducts two parallel analyses, one based on aggregate set-level data, and one based on detailed point-by-point data, which is available for a selected subsample of matches in three of the eight tournaments under examination.

The set-level analysis indicates that both men and women perform less well in the final and decisive set of the match. This result is robust to controls for the length of the match and to the inclusion of match and player-specific fixed effects. The drop in performance of women in the decisive set is slightly larger than that of men, but the difference is not statistically significant at conventional levels. On the other hand, the detailed point-by-point analysis reveals that, relative to men, women are substantially more likely to make unforced errors at crucial junctures of the match. Data on first serves suggests that women play a more conservative and less risky strategy as points become more important.

* PRELIMINARY AND INCOMPLETE. I thank Dan Ariely, Antonio Rangel, seminar participants at Hebrew University, 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. 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, 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. Former Harvard University President Lawrence Summers (2005) sparked a large controversy when he remarked that one of the reasons for the small number of women in science, engineering and at the forefront of academic research may lie in the difference in the distribution of "talent" between the two sexes. Even small gender differences in the standard deviation of talent can lead to very large differences in the number of men and women at the very high end of the distribution, where one is likely to find top corporate managers, leaders of the professions, or outstanding scientific researchers.

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

[1] A similar underrepresentation of women is also found among CEOs at Fortune 500 companies, tenured faculty at leading research institutions, conductors of philharmonic orchestras in the U.S., or top surgeons in New York City according to *New York* magazine.

performance does not. Gneezy et al. also found that women's lack of responsiveness to competitive incentives was restricted to mixed competitions (women competing against men), while in single-sex competitions (women competing against women) women did increase their performance under the competitive payment scheme. 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 (2005), on the other hand, found no gender differences in performance on an arithmetic task under either a non-competitive piece rate 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 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.

In a non-experimental setting, Lavy (2004) 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. Lavy attributes this finding to discrimination on the part of teachers, but one cannot rule out that women underperform because of the increased pressure they face in the blind setting.

The social psychology literature has studied at length the phenomenon of "choking under pressure", i.e., suboptimal performance despite a high degree of achievement motivation. 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), expectations of negative consequences (Paulus, 1983), and public expectations (Baumeister, Hamilton, and Tice, 1985; Strauss, 1997; and Strauss,1998).

The presence of a supportive audience (Butler and Baumeister, 1985), or the mere presence of others might also create pressure and induce individuals to choke (Zajonc, 1965). More recently, Ariely et al. (2004) also 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) complements the existing experimental evidence with 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 focus on professional tennis players and on all eight Grand Slam tournaments played in 2005 and 2006. One of the advantages of using sports data is that it is possible to observe detailed measures of productivity or performance. I ask whether players' performance deteriorates as the stakes become higher. A player's performance is measured primarily by looking at the number of unforced errors hit during each match. To the best of my knowledge, this is an improvement over previous research on tennis data, which limited itself to information on the number of games or sets won by the players (Sunde, 2003).

I conduct two parallel analyses, one based on aggregate set-level data, and one based on detailed point-by-point data, which is available for a selected subsample of matches in three of the eight tournaments under examination. The aggregate set-level analysis indicates that both men and women perform less well in the final and decisive set of the match. This result is robust to controls for the length of the match and to the inclusion of match and player-specific fixed effects. I find that the performance of women is slightly inferior to that of men, but the difference is not statistically significant at conventional levels.

The detailed point-by-point data, available for a subset of matches played at the 2006 French Open, Wimbledon and US Open tournaments, allows me to create a less coarse 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. Importance 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 while men's performance does not vary much depending on the importance of the point, women's performance deteriorates significantly as points become more important. About 30 percent of men's points end in unforced errors, regardless of their placement in the distribution of the importance variable. For women, on the other hand, about 34 percent of points in the bottom quartile of the importance distribution end in unforced errors; the percentage of unforced errors rises to nearly 40 percent for points in the top quartile of the importance distribution. These differences persist and are statistically significant even in a multinomial logit model that controls for players' ability, tournament, set number, and individual match fixed effects. I also document that women actually have a higher first serve percentage on important points, but the probability that the server wins the point does not increase with importance. I interpret this finding as evidence that women adopt a more conservative and less risky strategy on important points.

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 describes the data for the set-level analysis and Section 3 presents the results. Section 4 introduces the point-by-point data, and describes the construction of the importance variable. Section 5 presents the basic results of the multinomial logit analysis. Section 6 assesses the robustness of the basic results to different specifications of the model, and evaluates a number of explanations for the gender difference in performance, from differences in risk aversion to differences in anxiety. Section 7 concludes.

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.[2] 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 by the player. The definition of unforced errors is critical for the purposes of this paper. Leo Levin of IDS Sports, who has compiled statistics for all the major tennis tournaments, and who, according to *Tennis Magazine*, was responsible for introducing the concept of "unforced errors" in 1982, 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.[3] 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

[2] 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.
[3] From *http://www.tennis.com/yourgame/asktheeditors/asktheeditors.aspx?id=1432*.

watching tennis matches) who are recruited and trained specifically in advance of the tournament.[4]

2. Set level analysis: data and summary statistics

I collected data on all eight Grand Slam tournaments played in 2005 and 2006. The four Grand Slam tournaments (the Australian Open, the French Open, Wimbledon, and the US Open) are the most important and 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 January and September 2006 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. The web sites record detailed match statistics, broken down by set, for every match played in six of the eight tournaments; for the 2005 and 2006 U.S. Open, the full statistics are available for every match from the third round onwards, and for selected matches played in the first two rounds. I have information on the final score in the set, the number of points played in the set, the length of the set, and a number of statistics on the performance of both players, including the number of unforced errors, the number of winners, and the number of valid first serves. In addition, I also recorded the players' 52-week ranking at the beginning of the tournament. The weekly ranking takes into account all results obtained in professional and satellite tournaments over the past 52 weeks, and is the most widely used measure to assess players' relative abilities. Following Klaassen and Magnus (2001), I calculate from the rankings each player's *ability rating* as $Rating = 8 - \log_2(Rank)$. Klaassen and Magnus justify the use of this variable as a

[4] I do not have information on the identity of the stat-keepers, so a priori one cannot rule out the possibility that the observed gender difference in performance is 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 stat-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.

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 has the advantage of taking on higher values for better players, a desirable feature for a measure of ability.

The key indicators of performance are the percentage of unforced errors and the percentage of winners. Clearly a player who hits few unforced errors and many winners will win on average a high percentage of points. I prefer to focus primarily on the percentage of unforced errors, since a winning shot may be as much the result of a weak shot that preceded it as it is of outstanding play. It should be noted that both measures may be reflecting risk as much as quality: a more aggressive strategy will generally lead to more winners and forced opponent errors, but it may also generate more unforced errors, as the player attempts more risky shots. However, I find that the partial correlation (after having netted out own ability and opponent's ability) between unforced errors on one side and winners, forced opponent errors and first serves on the other is negative. This suggests that, on aggregate, all the variables are more a reflection of performance than of risk. Players who are tired and lack concentration and effort make more unforced errors, hit fewer winners, force fewer opponent errors, and have a lower first serve percentage.

Table 1 presents summary statistics for the main variables in the analysis, broken down by gender and tournament. I keep all matches with available detailed statistics, but I drop matches that were abandoned before the end because of one player's injury. This leaves me with a sample of 3,274 sets in 899 matches for men, and 2,053 sets in 892 matches for women. Note that the unit of observation is the set: in other words, the table reports the percentage of unforced errors, winners and forced errors hit *by both players* in the course of a set. This is to avoid the obvious dependence between the types of points hit by two players playing against each other.

The summary statistics show that there are important differences in the typology of the game, both across genders and across tournaments. These differences will probably not come as a surprise to even the casual tennis observer. Men hit on average fewer unforced errors than women, more winners and force more opponent errors. There are however large differences across tournaments: on the slow clay court surface of the French Open, it is much harder to hit winners and the percentage of unforced errors soars for both men and women. At the opposite extreme, the fast grass court surface of Wimbledon leads players to hit slightly more winners and to force more opponent errors, while unforced errors decrease. The hard court surfaces of the Australian Open and the U.S. Open stand somewhere in the middle. The percentage of first serves is similar across genders, even though there is some variation across tournaments, again with faster surfaces inducing higher first serve percentages for both men and women.

3. Set-level Analysis: Results

I first investigate whether performance decreases systematically as the stakes increase, i.e. as the match reaches its crucial stages. I start by using the individual set as the unit of analysis, and take the combined performance of both players together as the dependent variable. I adopt two indicators for the level of the stakes: 1) a dummy indicator for whether the set in question is the decisive set of the match, i.e., the third set in women's matches, and the fifth set in men's matches; 2) a continuous variable denoting the *importance* (or pivotality) of each set, defined as:

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

Klaassen and Magnus (2001) define in an analogous manner the importance of each point, and I will also use their definition in the point-level analysis that follows. This definition is simply the natural extension of their definition to set level data.[5] The probabilities are calculated assuming that the probabilities of winning a set are a fixed function of the two players' ability ratings. These probabilities are estimated using a probit model, separately for each gender and tournament. Note that the importance variable always takes on a value of 1 in the final and decisive set of a match, while it differs from the simple "final set" dummy in that it takes on values greater than zero in the earlier sets.

Letting $STAKES_{mt}$ be either the final set dummy or the importance variable in set t and match m, and stacking all the tournaments together, I run regressions of the following form, separately for men and women:

$$Performance_{mt} = \beta_0 + \beta_1 STAKES_{mt} + \gamma' X_{mt} + c_m + u_{mt},$$

where the performance variable represents the total percentage of shots of a given type played by *both* players (e.g., the percentage of unforced errors by both players out of the total number of points played in a set). The set of control variables includes a full set of tournament dummies interacted with the two players' rating, the tournament round, and the cumulative number of points played up to the beginning of the set, to capture any effects of fatigue on performance.[6] The fixed effect c_m captures any residual unobserved factor that may affect performance in match m, such as the weather, the physical health of a player, the way the two players match up against one another, and so on.

I estimate the equation both with and without match fixed effects. In the specification without fixed effects the identification of the coefficient of interest comes from variation in the performance variable both between and within players and matches. To account for

[5] It is a matter of straightforward algebra to show that based on this definition the importance of a set (or of a point) is the same for both players, so that it is appropriate to talk of "the" importance of the set (or point).

[6] Ideally, we would have wanted to control for the cumulative duration (in minutes) of the match up to the beginning of the set. Unfortunately, data on match duration are unavailable for one of the tournaments (Wimbledon 2005), so we proxy for duration with the length of the match in points. The correlation between match duration (in minutes) and match length is above 0.95 for both genders.

potential within-match correlation in the error terms, standard errors are adjusted for clustering at the match level. In the specification with match fixed effects, identification comes exclusively from variation in a player's performance *within* a match, i.e., whether the quality of the player's game deteriorates as the stakes of winning the set become larger. Clearly, in this specification, all the fixed player characteristics drop out of the equation, and only the final set dummy and the cumulative match length remain in the equation.

I first estimate the equation separately for the two genders, and I then stack all the data and add interactions between the female dummy and all the explanatory variables. In this stacked specification, the interaction between the final set indicator (or the importance variable) and the female dummy represents the difference between women and men in the effect of the final set (or importance).

The results are presented in Table 2. The top three rows present the results using the final set dummy as the measure of the magnitude of the stakes, while the bottom three rows use the importance variable defined above. Each row presents results for a different dependent variable. For both men and women, performance as measured by either unforced errors or winners deteriorates in the final set, regardless of whether we control for match fixed effects. On the other hand, the percentage of first serves does not change in the final set for either men or women. The percentage of unforced errors rises by between 1.14 percent and 1.36 percent for men, and by about 2.85 percent for women. For both men and women, the percentage of winners drops by 1.1-1.4 percent (men), and by 1.4-1.5 percent (women). These effects are not very large – the standard deviation in the percentage of unforced errors and the percentage of winners are about 12 and 8 percentage points respectively – but the effects are estimated fairly precisely. The drop in performance for women is always statistically significant at the 5 percent level, while the coefficients for men are statistically significant only in the fixed effects specification. The women's drop in performance is larger than that of men, but the differences are never statistically significant at conventional significance levels. Broadly similar results are obtained using the importance variable as the measure of the stakes: the only notable difference is that the drop in the percentage of winners is now larger for men than for women.

One possible explanation for the final set effect is that it simply reflects fatigue: as the match progresses and enters the final set, players are obviously more tired, and hence are more likely to make errors and less likely to hit winners. In Table 3 I present results from a series of additional regressions meant to address the fatigue hypothesis. The dependent variable in all the regressions is the percentage of unforced errors, and the stakes variable is the final set dummy. The odd-numbered columns simply replicate the results from Table 2, but now the cumulative points variable is entered linearly, rather than interacted with the tournament dummies. In all the specifications, the coefficient on the length of the match is large and *negative*, and highly statistically significant. The longer the match, the less likely are players to make unforced errors, in contrast to a simple fatigue-based explanation. In the even-numbered columns, I interact the final set dummy with cumulative points: this essentially asks whether the tendency to increase the number of unforced errors in the final set is stronger when the earlier sets had been longer. We find some evidence in support of this hypothesis for men, but none for women. This is not entirely surprising, given that men play on a "best-of-five" sets basis instead of a "best-of-three," and are therefore more likely to be fatigued and make mistakes when the final set comes along.

Overall, the results from the set-level analysis indicate that both men and women perform less well in high-stakes situations. Part of the explanation for men's drop in performance may be due to fatigue, but a fatigue-based explanation does not seem very plausible for women. It therefore appears possible that women cope less well with pressure, but most of the gender differences are small and not statistically significant. The analysis, however, is limited because of the coarseness of our measure of pressure. There are many critical junctions in a match that occur well before the final and decisive set. A break point in the latter stages of an evenly fought early set can be more decisive for the fate of a match than a point in the early stages of the final set. Therefore, I now move to the analysis of point-by-point data, where I will be able to construct a more refined measure of the level of the stakes at each stage of the match.

4. Point-by-point data: description and the importance variable

For the last three tournaments in the sample, the 2006 French Open, Wimbledon, and US Open, I was able to collect detailed point-by-point data for a selected number of matches that were played on the main championship courts and were covered by IBM's Point Tracker technology. For every single point played in these matches, I recorded who won the point, whether the first serve was in, who hit the last shot, the way the point ended (winner, unforced error, forced error, ace, double fault), and the score of the match. This data is available for a total of 187 matches, 102 for men and 85 for women. Altogether, I have data for more than 33,000 points that were played in these matches.

One of the key objectives of the analysis using point-by-point data is to construct a measure of the importance of each point. 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:

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

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. These fixed probabilities are calculated using the full 2005 French Open, Wimbledon and US Open data by running regressions of the proportion of points won by the server (the receiver) in each match as a function of the rating of the two players, separately by tournament and gender. 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 calculates recursively, for every pairing of players, the probability of winning the match at every possible stage. From this procedure it is then possible to calculate the importance of each point.

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.

Table 4 presents summary statistics for the importance variable, separately by tournament and gender. The mean of the importance variable is 0.0239 for men, and 0.0268 for women, reflecting the fact on average points in the women's game are more important, given that 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 5, 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 = 0.088) are on average about 5 times more important than points in the first set (average importance =

0.018). 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 0.0429, about twice as large as when the score is 0-0 (0.0230), and nearly 30 times as large as when the score is 5-0 (0.0014). There is also substantial variation within games: the average of the importance variable when the score is 40-0 is 0.0038, compared to 0.0493 when the score is 30-40.

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.

5. Point-by-point analysis: results

Table 6 presents the typology of points played by men and women, split by quartiles of the importance variable. The results are fairly striking. For men, there appears to be no systematic pattern in the typology of points by importance quartile: the percentage of unforced errors hovers between 29 and 30 percent, regardless of the importance of the point. About 55 percent of these unforced errors are made by the server, and the remainder by the receiver. On the other hand, the typology of points in women's matches is strongly affected by the importance variable. In the bottom quartile of the importance distribution, unforced errors are only slightly more frequent than winners or forced errors, and this distribution is not too dissimilar from that of men. However, as the importance of the points grows, women commit a growing number of unforced errors, with the percentage of winners and forced errors falling, especially the latter. In the top quartile of the importance distribution, the percentage of unforced errors reaches nearly 40 percent, more than five percentage points higher than what it was in the bottom quartile.

At a first glance, these results suggest that men and women react differently to increases in competitive pressure, with women exhibiting a lower level of performance as the stakes become higher. 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), or are more likely to be played at the French Open, where unforced errors are more frequent. To address these concerns, I proceed to a multinomial logit analysis for the typology of the point.

Specifically, define Y_{im} as the outcome of point i in match m. Y_{im} can take on three possible values: 1 – forced error, 2 – unforced error, 3 – winner. I estimate the following multinomial logit model:

$$P(Y_{im} = k) = \frac{\exp\left(\alpha_k + \beta_k IMP_{im} + \gamma_k' X_{im} + c_{mk}\right)}{1 + \sum_{k'=2}^{3} \exp\left(\alpha_{k'} + \beta_{k'} IMP_{im} + \gamma_{k'}' X_{ik'} + c_{mk}\right)}, \qquad k=2,3.$$

The main coefficients of interest are the β's, the coefficients on the importance variable, and in particular β_2, the effect of importance on the propensity to commit unforced errors. The set of control variables includes the server's ability rating, the receiver's ability rating, set dummies, the round of the tournament, the serial number of the point within the match to control for possible fatigue effects, a dummy indicating whether the match was played on the tournament's main court, and tournament dummies interacted with all of the preceding variables. The base category is forced errors: this implies that the β's represent the increase in the log odds of unforced errors (or winners) relative to forced errors when the importance variable increases by one unit. In addition to this basic specification, I also estimate two additional models (quadratic and piecewise constant) to detect potential nonlinearities in the effect of importance. The model is estimated with and without match fixed effects. In the model without match fixed effects, 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. Finally, standard errors are adjusted for clustering at the match level, to account for the fact that we have multiple observations coming from the same match.

The results of the multinomial logit analysis are shown in Table 7. The top panel presents the coefficients on the importance variables in the three models (linear, quadratic and piecewise constant) for the "unforced errors" equation, which are the main coefficients of interest, while the bottom panel presents the coefficient for the "winners" equation.

In the linear model, none of the coefficients on the importance variable are statistically significant, although the pattern of signs suggests that women are more likely to make unforced errors as importance grows. The quadratic equation reveals that there are some meaningful non-linearities in the effect of importance that were overlooked in the linear specification. For men, both the linear and the quadratic terms in the unforced error equation are fairly small and not always statistically significant. For women, both coefficients are significant at the 5 percent level, even after controlling for match fixed effects. The positive coefficient on the linear term and the negative coefficient on the quadratic term imply that the log odds of unforced errors relative to forced errors increases as importance grows, but at a declining pace. The turning point in the quadratic equation occurs at importance levels between 0.12 and 0.13, beyond the top 95th percentile of the importance distribution. The gender differences in the coefficients of the "unforced error" equation are statistically significant at the 10 percent level. The coefficients in the "winners" equation for women have a similar pattern, implying that the odds of a point ending in a winner also increase with importance. However, the coefficients are smaller and not always significant and the difference between men and women is not statistically significant.

The piecewise constant specification confirms and reinforces the results of the quadratic specification. In particular, we find that there is a significant difference between genders in the propensity to make unforced error in the top quartile of the importance distribution. This difference stems from an increased propensity to make errors by women in the no fixed-effects specification, and from a combination of decreased propensity by men and increased propensity by women when fixed match effects are controlled for. By contrast, there appears to be no significant relationship between the propensity to hit winners and the importance quartile for both genders. The size of the coefficient implies that the odds of women making unforced errors (relative to forced errors) rise by 19-26 log points when moving from the bottom to the top quartile of the importance distribution. To put this into perspective, the odds of making unforced errors rise by about 70-100 log points when moving from the fast courts of the US Open or Wimbledon to the slow clay courts of the French Open. Therefore, the impact of importance on women's propensity to make unforced errors is about one-fifth to one-third as large as the impact of the playing surface. Given the importance of the playing surface in determining tennis outcomes, this seems like a fairly large effect.

6. Extensions

Robustness checks

In Table 8 I assess the robustness of the results to a variety of different specifications. The coefficients in the previous table, coming from a multinomial logit model, reflect the propensity of making unforced errors *relative* to forced errors. One may view this as a fairly unusual measure of performance, and it may seem more appropriate to focus on the absolute propensity of making unforced errors, relative to any other type of shot. The top panel in Table 8 addresses this point, by estimating a simple logit model for the propensity of making unforced errors. The results are very much in line with those of Table 7: the odds of women making unforced errors rises by 18-19 log points when moving from the bottom to the top quartile of the importance distribution. The coefficient is statistically significant in the specification without fixed effects, and marginally insignificant in the specification with fixed effects, but the gender difference is always significant at the 10 percent level.

The next panel explores the sensitivity of the results to the definition of the importance variable. One may be worried that the results are sensitive to the calculation of the probabilities of winning a point, which were estimated on 2005 data using the players' ability ratings. To address this point, I calculated an alternative measure of importance, which assumes that the probability of winning a point does not depend at all on the players' ability, but rather is constant for each tournament, gender and serving status. The results of the analysis with the new importance variable are in the two bottom panels of Table 8: the multinomial logit model (analogous to that of Table 7) in panel B, and the logit model for the simple binary dependent variable in panel C. The effect of importance on women's unforced errors is still positive, but substantially smaller and no longer statistically significant in the multinomial logit model. It is possible that the results with the new measure of importance are attenuated, since we are introducing noise by ignoring that the amount of competitive pressure differs depending on the ability of the players. It is somewhat reassuring, though, that the general pattern of the coefficients does not change much with the new measure.

Figure 3 presents an alternative way of looking at the relationship between importance and the propensity to make unforced errors for the two genders. It eliminates the importance variable altogether, and focuses exclusively in variations in performance at different stages within the *game*. Specifically, I estimated a multinomial logit model as in Table 7, 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 percentage of unforced errors at every combination of points for a representative match, and plotted it against the average value of the importance variable at that combination. The figure shows that for men there is only a weak correlation between average importance of a point combination and the propensity to make unforced errors. For women, on the other hand, importance explains more than 40 percent of the variation in the propensity to make unforced errors within games.

Risk aversion

I now attempt to investigate more closely the possible determinants of the gender gap in the propensity to make errors at crucial stages of the match. One potential explanation is that women adopt a less aggressive and less risky strategy as points become more important. There is a large amount of evidence, both from the lab and from the field, that women are more risk averse than men (see Croson and Gneezy, 2004, for an overview). In the tennis setting, if both players adopted a safe hitting strategy (i.e., they just put the ball in play), they would be less likely to hit winners and force opponent errors, and naturally the rally would likely end in an unforced error. Ideally, to test this hypothesis we would need data on the strength and depth of each stroke, or on the length of each rally. This data is not easily available, but we can try to use information on first serves to gauge the degree of risk taking among men and women. In the top panel of Table 9, we analyze whether the importance of the point has any effect on the probability of the first serve being in. For men, importance does not matter at all in the specification without match fixed effects, and seems to have a small, but mostly insignificant effect in the specification with fixed effects. On the other hand, women are substantially more likely to hit the first serve in as the point becomes more important: the odds of hitting the first serve in increases by 18-25 log points when moving from the bottom to the top importance quartile. The difference between men and women is statistically significant in the specification without match fixed effects, and marginally significant at the 10 percent level in the specification with fixed effects. This result would suggest that women's performance actually increases as points become more important. However, hitting the first serve in is no guarantee of winning the point. In the bottom panel of Table 9, I look at whether the server is more likely to win the point as importance increases. Surprisingly, the higher percentage of first serves for women at important points does *not* translate into a higher percentage of points won by the server. Quite the contrary: women servers are slightly less likely to win points in the top quartile of the importance distribution than in the bottom quartile of the importance distribution (although the differences are not statistically significant). I interpret this finding as evidence that women hit "softer" first serves, which are more likely to be in, but that do not convey a significant advantage to the server. In other words, women adopt a more conservative playing strategy and take fewer risks. In a tournament setting such as this, this is not necessarily a good thing: it is not difficult to conjure up a model in which some degree of risk-taking is preferable to a completely safe strategy.

Further evidence that women may be hitting "softer" shots is provided in Table 10, which looks at the effect of importance on serve speed and on the length of rallies. These variables are available for about 90 percent of the matches in the point-by-point data set. Panel A of the table looks at the effect of the importance quartile on the speed of the first serve. Once again, we observe substantial gender differences in serve speed as the stakes become higher. Men hit slightly faster first serves as importance rises, but the effects are statistically significant only in the specification with fixed effects. On the other hand, women hit significantly slower first serves as the stakes become higher. This is true in both specifications, both with and without match fixed effects.[7] Panel B of the Table shows that the

[7] It should be noted that these estimates suffer from potential selection bias, since we observe the speed of the first serve only if the first serve is valid. However, it is straightforward to show that the likely direction of the

speed of women's second serve also declines markedly with importance: in the top importance quartile, the average speed of the second serve is nearly three and a half miles per hour slower than in the lowest importance quartile.

Finally, the bottom panel of the table looks at the effect of importance on the number of strokes in the rally. If women hit softer strokes, as it appears from the analysis of the speed of the serve, then they should also be involved in longer rallies. Then, one possible interpretation for the increased frequency of unforced errors at important stages of the match is that women play a less aggressive strategy, which unavoidably will result in an unforced error. The results show quite unambiguously that women play longer rallies as the importance of the point increases. The striking finding of the table, though, is that importance has an even stronger positive effect on the length of men's rallies. Hence, it appears that both men and women choose a more patient playing strategy at important stages of the match, but while men are able to sustain the rally and eventually build up the point and play effective strokes, women's rallies end more suddenly because of unforced errors.

Anticipation of winning and fear of losing

By definition, the importance variable is a measure of the magnitude of the stakes of each point, which is common to both players. However, it is conceivable that the type of pressure faced by a player who is close to winning the match is not the same as the type of pressure faced by the player who is close to losing. To address this point, we decompose the importance variable as follows:

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

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

$$\text{Prob (1 wins match | 1 wins point } t\text{) - Prob (1 wins match | current status) +}$$

$$\text{Prob (1 wins match | current status) - Prob (1 wins match | 1 loses point } t\text{) =}$$

---

bias is positive, meaning that the simple OLS estimates reported in the table are probably even smaller (in absolute value) than those that would have resulted from taking into account sample selection.

$$Anticipation_t + Fear_t.$$

We call the first term in the sum the *anticipation of winning*: it measures by how much the probability of winning the match increases if the player wins the current point, relative to the probability of winning the match before the point is played. Put more simply, it indicates how much winning the current point brings the player closer to victory in the whole match. The second term is the *fear of losing*. It tells us by how much the probability of winning the match decreases if the player loses the current point, or by how much losing the current point brings one closer to defeat. It is straightforward to show that for any given point, the anticipation of winning for one player is equal to the fear of losing for the other. To analyze whether anticipation and fear affect performance, I estimate multinomial logit models where the dependent variable has now six possible outcomes: winner by the server, winner by the receiver, unforced error by the server, unforced error by the receiver, forced error by the server and forced error by the receiver. I estimate the model twice, once using four quartiles of *fear* as the explanatory variable, and once using four quartiles of *anticipation*. The results are presented in Table 11: to avoid cluttering the table with too many numbers, I present only the effects of the fourth quartile of the fear/anticipation variable on the odds of making unforced errors relative to forced errors, for the specific player under investigation. The results show that, with one exception, the fear of losing and the anticipation of winning affect women's propensity to make unforced errors in roughly equal measures. Being in the fourth quartile of the fear distribution instead of the first raises the odds of making unforced errors by 19-31 log points, while being in the top quartile of the anticipation distribution raise the odds by 1-30 log points. Overall, it is difficult to make a strong statement about whether it is the fear of losing or the anticipation of losing that induces a higher percentage of unforced errors.

<u>Individual determinants of reduced performance</u>

Finally, we may be interested in identifying which individual characteristics are significant determinants of reduced performance under pressure. To do this, we adopt the following strategy: we first reshape the data so that every observation is a player-point combination (instead of a single point). We then estimate, separately by gender, a linear probability model for unforced errors, and include, in addition to the usual control variables, a full set of player-match dummies, and the player-match dummies interacted linearly with the importance variable. These player-match-specific slopes represent the individual propensity of each player to make unforced errors in a given match. We then regress these coefficients on the own player's ability rating, on the opponent's ability rating, and on a number of physical characteristics of the player (age, height, and weight). The results are presented in Table 12. The odd-numbered columns are simple unweighted regressions, while the even-numbered columns use as weights the number of observations per player in the first-stage regression. For both men and women we find that the propensity to make unforced errors as points become more important decreases with the player's own ability rating, and increases with the opponent's rating, but the effects are small and not always statistically significant. As for the player's physical characteristics, they add only little explanatory power to the regressions. There is some weak evidence that for men the propensity to underperform in pressure situations increases with age, and then decreases, while no such effect is observed for women. This is at odds with the hypothesis that more experienced players may have learnt to cope better with pressure situations. Overall, it appears that a large fraction of the observed variability in the propensity to make unforced errors at crucial stages of the match remains unexplained.

7. Conclusion

In this paper I have used data from the eight Grand Slam tournaments played in 2005 and 2006 to assess whether men and women respond differently to competitive pressure in a real-world setting with large monetary rewards. The aggregate set-level data reveals that the performance of both men and women deteriorates in the final and decisive set. Women's decline in performance is more pronounced than that of men, but the difference is not statistically significant. On the other hand, the analysis using detailed point-by-point data indicates that there are significant differences between men's and women's performances at crucial junctions of the match: the propensity of women to commit unforced errors increases significantly with the importance of the point, while men's propensity to commit unforced errors is unaffected by point importance. It appears that some of this difference can be explained by gender differences in risk-aversion: the evidence on first serves and on the percentage of points won by the server strongly suggests that women tend to adopt a safer and less aggressive strategy on important points. This itself can be a mark of low performance: in a tournament setting, some degree of risk-taking may be desirable.

To what extent then can we draw from this study more general lessons about gender differences in the labor market? An unforced error is by definition an error that cannot be attributed to any factor other than poor judgment by the player. Can we extrapolate from our findings that in general women's judgment becomes more clouded as the stakes become higher, and this may hinder their advancement to the upper echelons of management, science, and the professions? Clearly, the answer must be negative. The results are only relevant for the specific context, and it is questionable whether the conclusions can be even extended 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, there are at least two striking features in this study that still deserve attention. First, the women in our sample are among the very best in the world in their profession, and are without question extremely competitive. They are probably quite distant from the typical woman in experimental studies, which underperforms in competitive settings and shies away from competition. Therefore, it is doubly surprising that even these highly competitive women exhibit a decline in performance in high pressure situations. In many respects, this sample is more representative of the extreme right tail of the talent distribution that is of interest for understanding the large under-representation of women in top corporate jobs, prestigious professions and academia. Second, some experimental studies (e.g., Gneezy, Niederle, and Rustichini, 2003) found that women's tendency to underperform in competitive environments occurs only when they compete against men. By contrast, here we find that women's performance deteriorates as competitive pressure rises, even when the competition is clearly restricted to women alone. This may have implications for educational policies such as single-sex schooling, and deserves further investigation.

References

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.

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

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), pp. 1213-1230.

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

Dohmen, Thomas 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*, August 2003.

Gneezy, Uri and Rustichini, Aldo. "Gender and Competition at a Young Age". *American Economic Review,* May 2004.

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*, June 2001.

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*, 1988, 1, pp. 235-246.

Lavy, Victor. "Do Gender Stereotypes Reduce Girls' Human Capital Outcomes? Evidence from a Natural Experiment." NBER Working Paper No. 10678, August 2004.

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?" NBER Working Paper No. 11474, June 2005.

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.

Table 1: Summary Statistics

	Men					Women				
	Total	Australian Open	French Open	Wimbledon	US Open	**Total**	Australian Open	French Open	Wimbledon	US Open
Number of matches	899	253	249	244	153	892	252	254	245	141
Number of sets	3,274	925	900	876	573	2,053	578	586	570	319
Pct. unforced errors	32.72 (11.94)	33.95 (9.39)	41.96 (11.33)	22.22 (7.88)	32.30 (8.46)	40.48 (12.31)	44.73 (9.51)	47.45 (11.73)	30.11 (8.94)	38.48 (9.89)
Pct. winners	34.24 (7.95)	32.85 (8.24)	34.12 (7.50)	35.54 (7.81)	34.72 (8.03)	30.05 (7.98)	27.71 (7.86)	31.43 (7.80)	30.89 (7.29)	30.27 (8.76)
Pct. forced errors	33.03 (11.04)	33.20 (7.93)	23.92 (10.21)	42.24 (8.37)	32.99 (8.00)	29.47 (11.34)	27.56 (7.77)	21.11 (10.57)	38.99 (9.08)	31.25 (8.18)
Pct. first serve	60.68 (10.86)	58.88 (10.60)	61.59 (11.28)	62.25 (10.36)	59.76 (10.87)	61.64 (11.38)	59.52 (11.00)	61.94 (11.61)	63.09 (11.42)	62.31 (11.04)
Average player rank	84.07	84.80	83.86	87.31	78.04	79.17	80.15	80.96	82.74	68.97
Average player rating	2.22	2.16	2.19	2.15	2.47	2.27	2.21	2.21	2.18	2.67

Note: Data refers to all completed matches for which detailed statistics are available. Standard deviations in parentheses.

Table 2: The Effect of the Magnitude of the Stakes on Performance

Set-level analysis

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
A: Final set dummy						
Pct. unforced Errors	1.1392 [1.50]	2.8524 [3.43]	1.7132 [1.52]	1.3576 [1.86]	2.8542 [3.54]	1.4966 [1.38]
Pct. Winners	-1.1219 [-1.74]	-1.4900 [-2.29]	-0.3681 [-0.40]	-1.3943 [-2.22]	-1.3590 [-2.10]	0.0354 [0.04]
Pct. first Serve	-0.3130 [-0.50]	-0.0171 [-0.03]	0.3302 [0.36]	0.2637 [0.43]	-0.4544 [-0.68]	-0.7181 [-0.80]
B: Importance variable						
Pct. unforced errors	1.3383 [1.39]	2.4225 [1.95]	1.0842 [0.69]	0.7999 [0.84]	2.3893 [1.93]	1.5895 [1.02]
Pct. winners	-1.6806 [-1.90]	-1.1278 [-1.08]	0.4808 [0.36]	-2.2521 [-2.68]	-1.0068 [-1.02]	1.2452 [0.96]
Pct. first serve	-0.5962 [-0.75]	1.7874 [1.85]	2.3836 [1.25]	0.1506 [0.18]	0.2753 [0.26]	0.1247 [0.09]
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of observations	3274	2053	5327	3274	2053	5327

Note: Entries in the table represent the coefficient on the "decisive set" dummy, t-statistics in parentheses. All regressions control for a full set of tournament dummies interacted with: a constant, the high-ranked player's rating, the low-ranked player's rating, the cumulative number of points played up to the beginning of the set, and the tournament round. Standard errors are corrected for clustering at the match level.

Table 3: The Effect of the Length of the Game and the Decisive Set on the Percentage of Unforced Errors

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Final Set	1.141 (1.50)	1.230 (1.64)	1.3337 (1.92)	1.384 (1.92)	2.826 (3.39)	2.924 (3.49)	2.830 (3.51)	2.802 (3.46)
Total number of points played before start of set	-0.010 (-4.49)	-0.011 (-4.71)	-0.010 (-4.61)	-0.010 (-4.73)	-0.0235 (-3.97)	-0.024 (-4.11)	-0.027 (-4.77)	-0.027 (-4.63)
Final Set × number of points (centered)	-	0.050 (2.17)	-	0.049 (2.33)	-	0.0147 (0.51)	-	-0.008 (-0.29)
Match fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Number of observations	3,274	3,274	3,274	3,274	2,053	2,053	2,053	2,053

Note: Entries in the table represent the coefficient on the relevant variables, t-statistics in parentheses. The interaction term is centered at the mean number of points at the beginning of the set. All regressions control for a full set of tournament dummies interacted with: a constant, the high-ranked player's rating, the low-ranked player's rating, and the tournament round. Standard errors are corrected for clustering at the match level.

Table 4: The importance of points – summary statistics

	Men				Women				
	French Open 2006	Wimbledon 2006	US Open 2006	All	French Open 2006	Wimbledon 2006	US Open 2006	All	**ALL**
Mean	0.0220	0.0243	0.0255	0.0239	0.0316	0.0236	0.0226	0.0268	**0.0248**
Standard deviation	0.0299	0.0329	0.0373	0.0335	0.0411	0.0343	0.0363	0.0383	**0.0352**
25^{th} percentile	0.0023	0.0030	0.0028	0.0027	0.0052	0.0013	0.0004	0.0020	**0.0025**
50^{th} percentile	0.0116	0.0129	0.0133	0.0125	0.0170	0.0092	0.0060	0.0118	**0.0123**
75^{th} percentile	0.0305	0.0322	0.0338	0.0321	0.0440	0.0314	0.0303	0.0374	**0.0336**
Number of points	7,776	7,076	7,529	22,381	4,728	2,587	3,532	10,847	**33,228**

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 5: 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 points	All	Men	Women
2-0	.0110	.0110	-	5-0	.0014	40-0	.0038	.0026	.0070
0-0	.0189	.0180	.0200	0-5	.0017	30-0	.0096	.0079	.0142
1-0	.0214	.0194	.0241	1-5	.0048	40-15	.0102	.0081	.0154
2-1	.0327	.0327	-	0-4	.0061	15-0	.0157	.0138	.0200
1-1	.0474	.0361	.0661	4-0	.0077	0-40	.0182	.0241	.0108
2-2	.0884	.0884	-	5-2	.0088	30-15	.0197	.0174	.0248
				2-5	.0088	0-0	.0197	.0185	.0223
				3-5	.0103	40-30	.0234	.0201	.0302
				5-1	.0106	15-15	.0257	.0244	.0281
				1-4	.0108	0-15	.0263	.0266	.0258
				0-3	.0118	0-30	.0278	.0315	.0221
				1-3	.0132	15-40	.0313	.0361	.0245
				2-4	.0134	deuce	.0334	.0323	.0354
				4-1	.0151	15-30	.0352	.0366	.0327
				3-0	.0155	30-40	.0494	.0529	.0438
				0-2	.0164	TB	.0743	.0741	.0752
				2-0	.0191				
				1-0	.0196				
				0-1	.0202				
				2-1	.0210				
				0-0	.0230				
				1-2	.0237				
				3-1	.0238				
				1-1	.0243				
				2-2	.0255				
				2-3	.0255				
				3-2	.0283				
				3-3	.0304				
				5-3	.0340				
				3-4	.0356				
				5-4	.0374				
				4-4	.0380				
				6-5	.0401				
				5-6	.0425				
				5-5	.0429				
				4-5	.0447				
				TB	.0743				
				6-6	.2269				

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 6: Typology of points, by importance quartiles

		Men				Women			
		Number of points	Pct. winners	Pct. unforced	Pct. forced errors	Number of points	Pct. winners	Pct. unforced	Pct. forced errors
Importance quartile 1 (least important)		**5364**	**30.98**	**30.54**	**34.12**	**2931**	**29.21**	**34.25**	**31.12**
	Server		21.32	17.19	8.31		16.68	19.65	9.62
	Receiver		9.66	13.35	25.80		12.52	14.60	21.49
Importance quartile 2		**5725**	**31.91**	**29.75**	**35.13**	**2569**	**29.86**	**37.06**	**30.67**
	Server		22.03	17.22	9.08		17.40	21.45	10.51
	Receiver		9.89	12.52	26.04		12.46	15.61	20.16
Importance quartile 3		**5944**	**30.52**	**30.79**	**35.41**	**2350**	**29.36**	**38.68**	**30.47**
	Server		21.35	17.78	9.24		17.32	22.30	9.45
	Receiver		9.17	13.00	26.18		12.04	16.38	21.02
Importance quartile 4 (most important)		**5312**	**31.23**	**30.61**	**34.73**	**2982**	**29.98**	**39.67**	**27.83**
	Server		21.44	17.43	9.09		18.04	22.07	8.38
	Receiver		9.79	13.18	25.64		11.94	17.61	19.45

Note: The importance quartiles are based on the overall distribution of the importance variable.

Table 7: The Effect of Importance on Performance
Multinomial Logistic Regression

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
Unforced Errors						
A: Linear						
Importance	-0.0230	1.3601	1.3831	-0.6381	0.6083	1.2464
	[-0.03]	[1.23]	[1.03]	[-0.85]	[0.49]	[0.86]
B: Quadratic						
Importance	1.3495	5.3462	3.9967	0.0789	4.3131	4.2342
	[1.24]	[2.98]	[1.91]	[0.07]	[2.13]	[1.82]
Importance	-7.3856	-21.8252	-14.4396	-3.3851	-16.5576	-13.1726
squared	[-2.19]	[-2.86]	[-1.73]	[-1.02]	[-2.06]	[-1.52]
C: Piecewise constant						
Importance	-0.0069	0.0736	0.0805	-0.1166	0.0764	0.193
quartile 2	[-0.12]	[0.99]	[0.87]	[-1.88]	[0.87]	[1.80]
Importance	0.0258	0.1521	0.1264	-0.1318	0.06	0.1918
quartile 3	[0.36]	[1.89]	[1.17]	[-1.81]	[0.50]	[1.37]
Importance	0.0182	0.2576	0.2394	-0.1557	0.1891	0.3448
quartile 4	[0.26]	[2.95]	[2.14]	[-2.03]	[1.45]	[2.28]
Winners						
A: Linear						
Importance	-0.8270	0.2205	1.0475	-1.4024	-1.2065	0.1959
	[-1.36]	[0.27]	[1.02]	[-2.42]	[-1.18]	[0.17]
B: Quadratic						
Importance	0.6286	3.3221	2.6935	-0.4505	1.4649	1.9154
	[0.71]	[2.25]	[1.57]	[-0.47]	[0.82]	[0.95]
Importance	-8.2234	-16.8999	-8.6765	-4.9305	-11.7872	-6.8567
squared	[-2.44]	[-3.05]	[-1.34]	[-1.43]	[-1.73]	[-0.90]
C: Piecewise constant						
Importance	0.0307	0.0057	-0.025	-0.0237	0.0361	0.0598
quartile 2	[0.60]	[0.07]	[-0.27]	[-0.39]	[0.43]	[0.58]
Importance	-0.0146	0.0479	0.0624	-0.1095	-0.0225	0.087
quartile 3	[-0.25]	[0.53]	[0.58]	[-1.53]	[-0.18]	[0.61]
Importance	-0.0033	0.1216	0.1248	-0.1124	0.005	0.1173
quartile 4	[-0.06]	[1.41]	[1.23]	[-1.65]	[0.04]	[0.80]
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of observations	22,345	10,832	33,177	22,345	10,832	33,177

Note: Entries in the table are the coefficients in a multinomial logit model for the typology of points (winners/ unforced errors/ forced errors) on the importance variables. Additional control variables: set dummies, server's rating, receiver's rating, serial number of the point within the match, tournament round, whether the match was played on center court, and interactions of all of the above with Wimbledon 2006 and US Open 2006 tournament dummies. The base category is "forced errors." Robust z-statistics (adjusted for clustering at the match level) in parentheses.

Table 8: Robustness Checks

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
A: Binary dependent variable (unforced error), logit model						
Importance	-0.0211	0.0705	0.0917	-0.1063	0.0582	0.1645
quartile 2	[-0.45]	[1.03]	[1.11]	[-2.05]	[0.67]	[1.63]
Importance	0.0321	0.127	0.0949	-0.0824	0.0694	0.1518
quartile 3	[0.52]	[1.68]	[0.98]	[-1.33]	[0.65]	[1.24]
Importance	0.0192	0.1939	0.1747	-0.1047	0.1837	0.2883
quartile 4	[0.30]	[2.64]	[1.81]	[-1.53]	[1.63]	[2.20]
B: Importance not a function of players' ability rating: Multinomial logit coefficients on unforced errors (base category: forced errors).						
Importance	0.0082	-0.0029	-0.0111	-0.0008	0.0308	0.0317
quartile 2	[0.17]	[-0.03]	[-0.12]	[-0.02]	[0.35]	[0.32]
Importance	-0.0632	-0.0215	0.0417	-0.0633	0.0022	0.0656
quartile 3	[-1.21]	[-0.24]	[0.41]	[-1.25]	[0.02]	[0.61]
Importance	-0.0457	0.097	0.1427	-0.036	0.1118	0.1478
quartile 4	[-0.77]	[1.03]	[1.29]	[-0.61]	[1.11]	[1.27]
C: Importance not a function of players' ability rating: Logit coefficients on binary dependent variable: unforced errors.						
Importance	0.0162	0.0425	0.0263	0.0138	0.0783	0.0645
quartile 2	[0.38]	[0.58]	[0.31]	[0.32]	[0.97]	[0.71]
Importance	-0.0344	0.0562	0.0906	-0.0257	0.0743	0.1000
quartile 3	[-0.76]	[0.76]	[1.05]	[-0.58]	[0.88]	[1.05]
Importance	0.0151	0.1521	0.1370	0.0320	0.1628	0.1308
quartile 4	[0.28]	[1.99]	[1.47]	[0.60]	[1.84]	[1.27]
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of observations (points)	22,345	10,832	33,177	22,345	10,832	33,177

Note: Entries in the table are the coefficients in either a logit model (panel A) or a multinomial logit model (panel B) on the relevant explanatory variables. Additional control variables: set dummies, server's rating, receiver's rating, serial number of the point within the match, tournament round, whether the match was played on center court, and interactions of all of the above with Wimbledon 2006 and US Open 2006 tournament dummies. In the multinomial logit models, the base category is "forced errors." Robust z-statistics (adjusted for clustering at the match level) in parentheses.

Table 9: The Effect of Importance on First Serve Percentage and on the Probability of the Server Winning the Point
Logistic Regression

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
A: Dependent variable: first serve in						
Importance quartile 2	0.0048	0.0906	0.0858	0.0953	0.1000	0.0047
	[0.10]	[1.66]	[1.20]	[1.89]	[1.44]	[0.05]
Importance quartile 3	-0.0255	0.1689	0.1944	0.0662	0.2185	0.1523
	[-0.52]	[2.49]	[2.33]	[1.22]	[2.47]	[1.47]
Importance quartile 4	-0.0196	0.1837	0.2032	0.0703	0.2456	0.1752
	[-0.35]	[3.34]	[2.59]	[1.20]	[2.76]	[1.65]
B: Dependent variable: Server wins point						
Importance quartile 2	-0.0462	-0.0672	-0.021	-0.036	-0.0963	-0.0603
	[-1.12]	[-1.16]	[-0.30]	[-0.72]	[-1.32]	[-0.68]
Importance quartile 3	-0.0481	-0.0204	0.0277	-0.0543	-0.0849	-0.0306
	[-1.08]	[-0.32]	[0.36]	[-0.91]	[-0.97]	[-0.29]
Importance quartile 4	-0.0349	-0.0074	0.0275	-0.0622	-0.1105	-0.0483
	[-0.61]	[-0.12]	[0.33]	[-0.91]	[-1.25]	[-0.43]
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of observations (points)	22,345	10,832	33,177	22,345	10,832	33,177

Note: Entries in the table are the coefficients on the importance variables in a logit model for whether the first serve is in (panel A), or whether the server won the point (panel B). Additional control variables: set dummies, server's rating, receiver's rating, serial number of the point within the match, tournament round, whether the match was played on center court, and interactions of all of the above with Wimbledon 2006 and US Open 2006 tournament dummies. Robust z-statistics (adjusted for clustering at the match level) in parentheses.

Table 10: The Effect of Importance on Serve Speed and on the Number of Strokes per Rally

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
A: Dependent variable: first serve speed (mph)						
Importance quartile 2	0.8662 [1.33]	-0.9469 [-1.56]	-1.8131 [-2.05]	0.7599 [1.91]	-0.5980 [-1.52]	-1.3579 [-2.44]
Importance quartile 3	0.4674 [0.59]	-1.6661 [-2.12]	-2.1335 [-1.91]	1.1764 [2.53]	-1.0831 [-2.19]	-2.2595 [-3.34]
Importance quartile 4	0.7818 [0.89]	-2.7670 [-3.35]	-3.5488 [-2.95]	1.5475 [3.06]	-1.4915 [-2.52]	-3.0389 [-3.91]
Number of observations	11,844	5,604	17,448	11,844	5,604	17,448
B: Dependent variable: second serve speed (mph)						
Importance quartile 2	0.0817 [0.13]	-1.2930 [-1.71]	-1.3747 [-1.39]	-0.7279 [-2.06]	-1.5774 [-3.31]	-0.8495 [-1.44]
Importance quartile 3	-0.2645 [-0.36]	-2.1540 [-2.24]	-1.8895 [-1.57]	-1.1304 [-2.55]	-2.5777 [-4.40]	-1.4473 [-1.98]
Importance quartile 4	-0.7540 [-0.99]	-2.1329 [-2.11]	-1.3789 [-1.09]	-1.6733 [-3.52]	-3.4714 [-5.12]	-1.7981 [-2.18]
Number of observations	6,679	2,947	9,626	6,679	2,947	9,626
C: Dependent variable: number of strokes per rally						
Importance quartile 2	0.1172 [0.96]	0.1263 [0.86]	0.0090 [0.05]	0.4416 [4.62]	0.4401 [3.05]	-0.0015 [-0.01]
Importance quartile 3	0.4496 [3.04]	0.1087 [0.72]	-0.3410 [-1.62]	0.9389 [8.55]	0.5726 [3.19]	-0.3663 [-1.75]
Importance quartile 4	0.8167 [4.65]	0.6027 [3.30]	-0.2140 [-0.85]	1.3960 [10.19]	1.1645 [6.27]	-0.2315 [-1.01]
Number of observations	19,504	9,047	28,551	19,504	9,047	28,551
Match fixed effects	No	No	No	Yes	Yes	Yes

Note: Entries in the table are the coefficients on the importance variables in a linear regression model. Additional control variables: set dummies, server's rating, receiver's rating, serial number of the point within the match, tournament round, whether the match was played on center court, and interactions of all of the above with Wimbledon 2006 and US Open 2006 tournament dummies. Robust t-statistics (adjusted for clustering at the match level) in parentheses.

Table 11: Anticipation of winning and fear of losing

	Men	Women	Difference	Men	Women	Difference
Fear of Losing						
Unforced errors, server						
Fear quartile 4	-0.0198	0.2224	0.2422	-0.0348	0.1870	0.2218
	[-0.21]	[1.62]	[1.46]	[-0.26]	[1.06]	[1.01]
Unforced errors, receiver						
Fear quartile 4	-0.0762	0.3080	0.3842	-0.2766	0.2201	0.4967
	[-0.77]	[2.70]	[2.55]	[-2.84]	[1.29]	[2.54]
Anticipation of Winning						
Unforced errors, server						
Anticipation quartile 4	-0.0408	0.2981	0.3389	-0.0727	0.2656	0.3383
	[-0.42]	[2.24]	[2.06]	[-0.49]	[1.37]	[1.39]
Unforced errors, receiver						
Anticipation quartile 4	-0.0403	0.2435	0.2476	-0.2228	0.0138	0.2366
	[-0.40]	[2.02]	[1.60]	[-1.97]	[0.07]	[1.07]
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of observations (points)	22,346	10,833	33,179	22,346	10,833	33,179

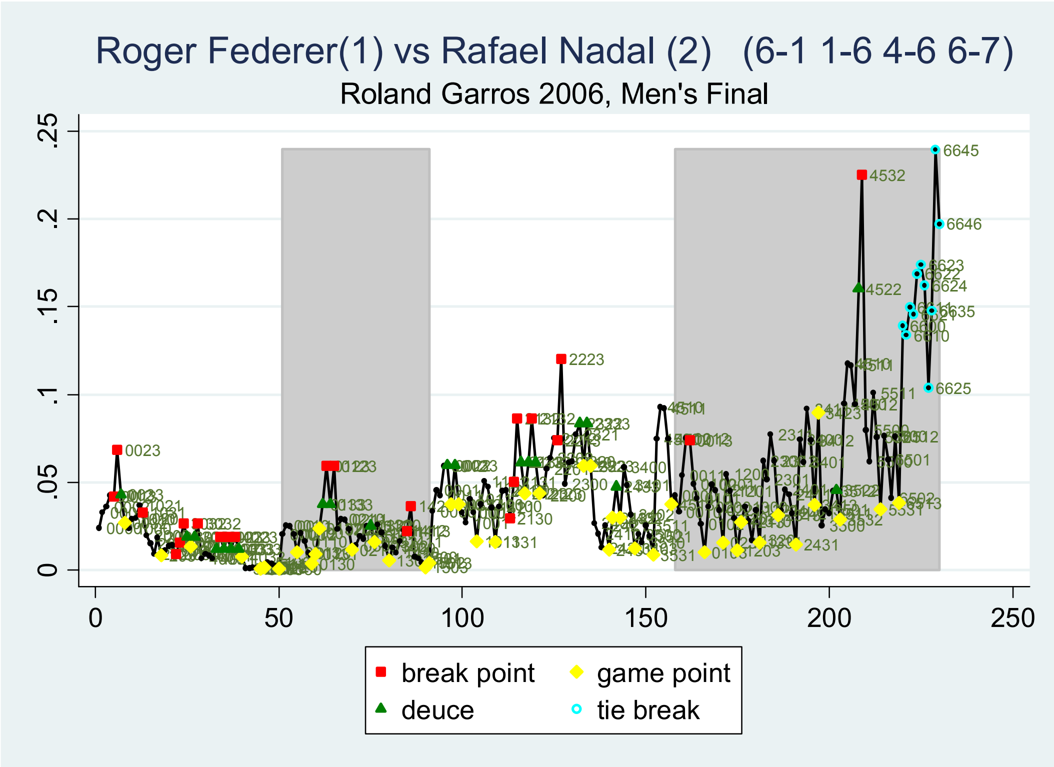
Note: Entries in the table represent the increase in log odds of unforced errors relative to forced errors for the server or the receiver, when moving from the bottom to the top quartile of the fear/anticipation distribution. Additional control variables: set dummies, server's rating, receiver's rating, serial number of the point within the match, tournament round, whether the match was played on center court, and interactions of all of the above with Wimbledon 2006 and US Open 2006 tournament dummies. Robust z-statistics (adjusted for clustering at the match level) in parentheses.

Table 12: Player characteristics and the propensity to make unforced errors at important points

Dependent variable: coefficients on player-match fixed effects interacted with importance in linear probability model of the probability of making unforced errors.

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Own Rating	-0.067 (-1.20)	-0.040 (-1.13)	-0.087 (-1.49)	-0.052 (-1.53)	-0.098 (-1.23)	-0.121 (-2.28)	-0.084 (-0.98)	-0.117 (-2.00)
Opponent's rating	0.133 (2.24)	0.144 (3.95)	0.124 (2.20)	0.138 (3.94)	0.169 (2.19)	0.117 (2.03)	0.161 (2.07)	0.109 (1.88)
Age	-	-	0.415 (1.80)	0.307 (1.79)	-	-	0.229 (0.62)	-0.066 (-0.20)
Age squared	-	-	-0.007 (-1.69)	-0.005 (1.65)	-	-	-0.003 (-0.35)	0.002 (0.31)
Height in cm	-	-	0.000 (0.00)	0.007 (0.36)	-	-	0.004 (0.21)	0.006 (0.34)
Weight in kg	-	-	-0.018 (-0.99)	-0.022 (-1.43)	-	-	-0.007 (-0.30)	-0.013 (-0.59)
weights	no	yes	no	yes	no	yes	no	yes
Number of observations	204	204	199	199	170	170	161	161
R^2	0.075	0.092	0.107	0.125	0.052	0.058	0.067	0.063

Note: OLS and WLS coefficients**,** t-statistics in parentheses.



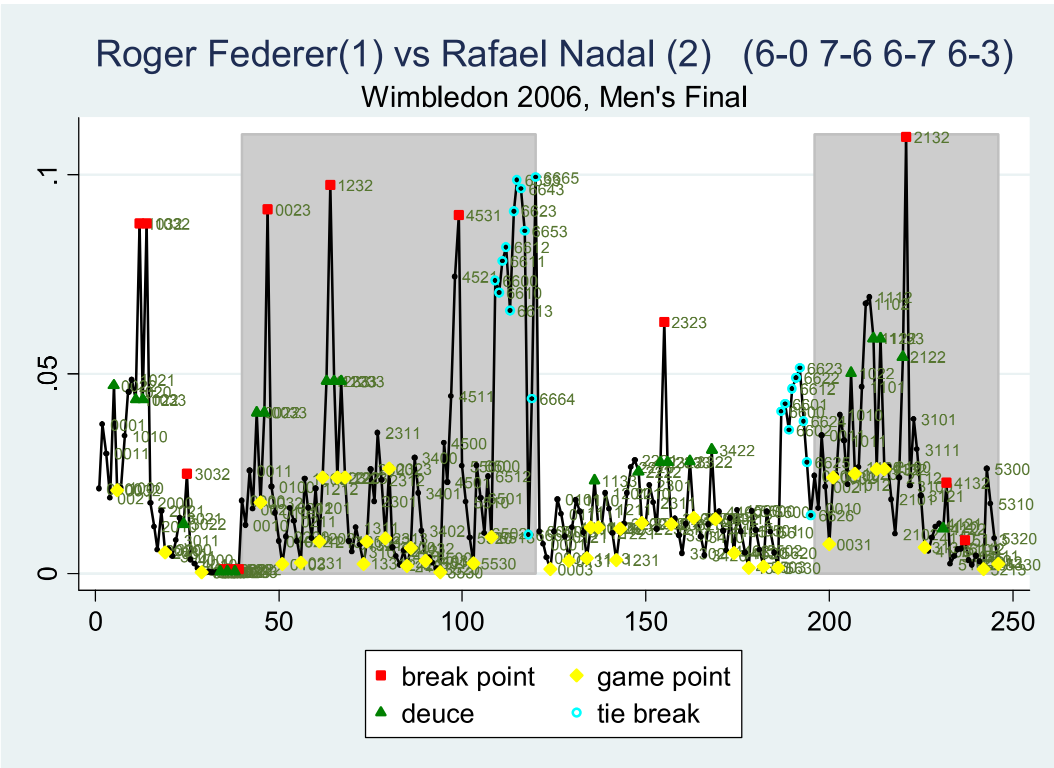
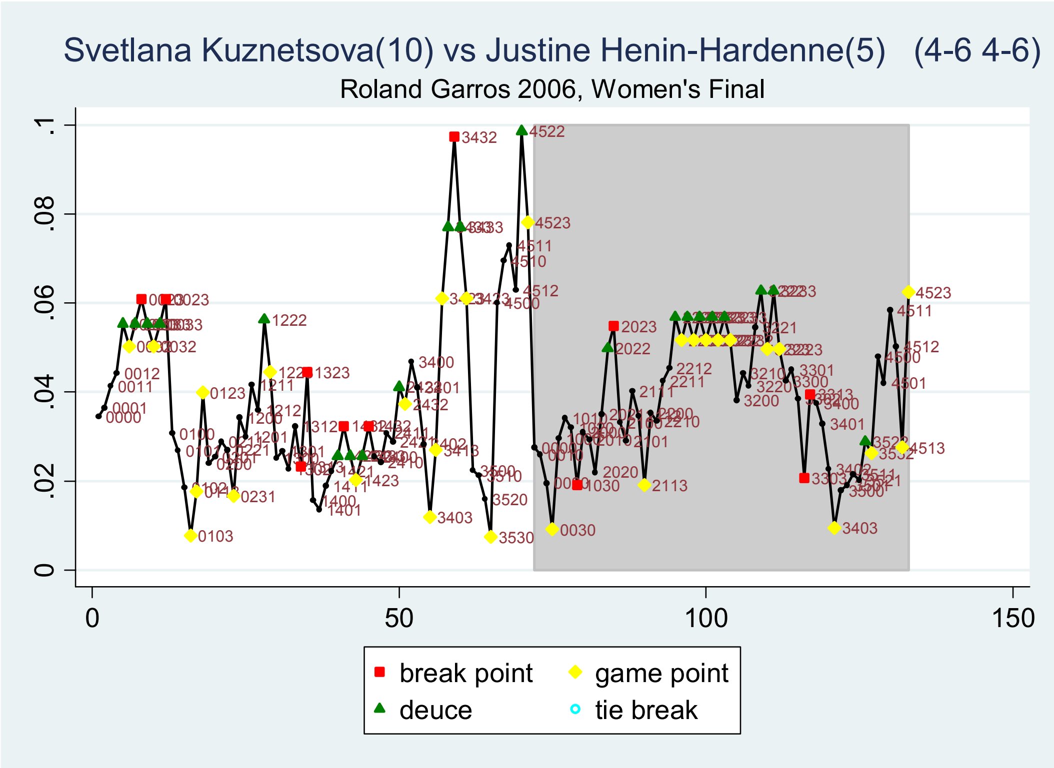


Figure 1: The evolution of importance over the course of selected matches – men



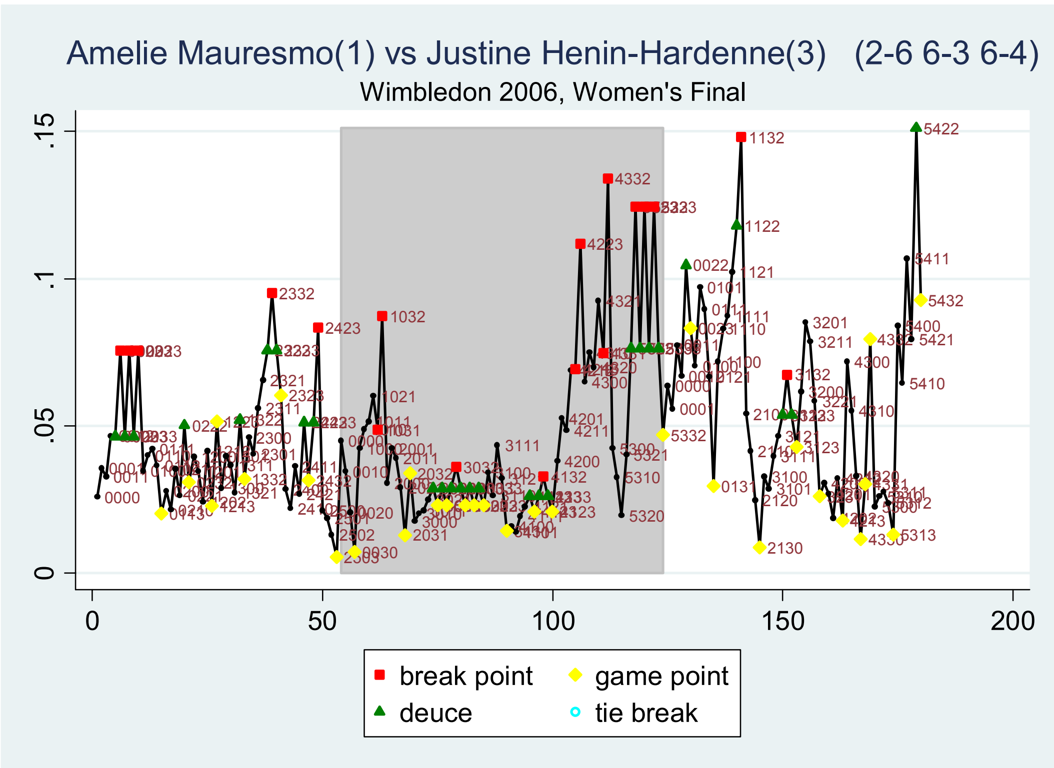


Figure 2: The evolution of importance over the course of selected matches – women

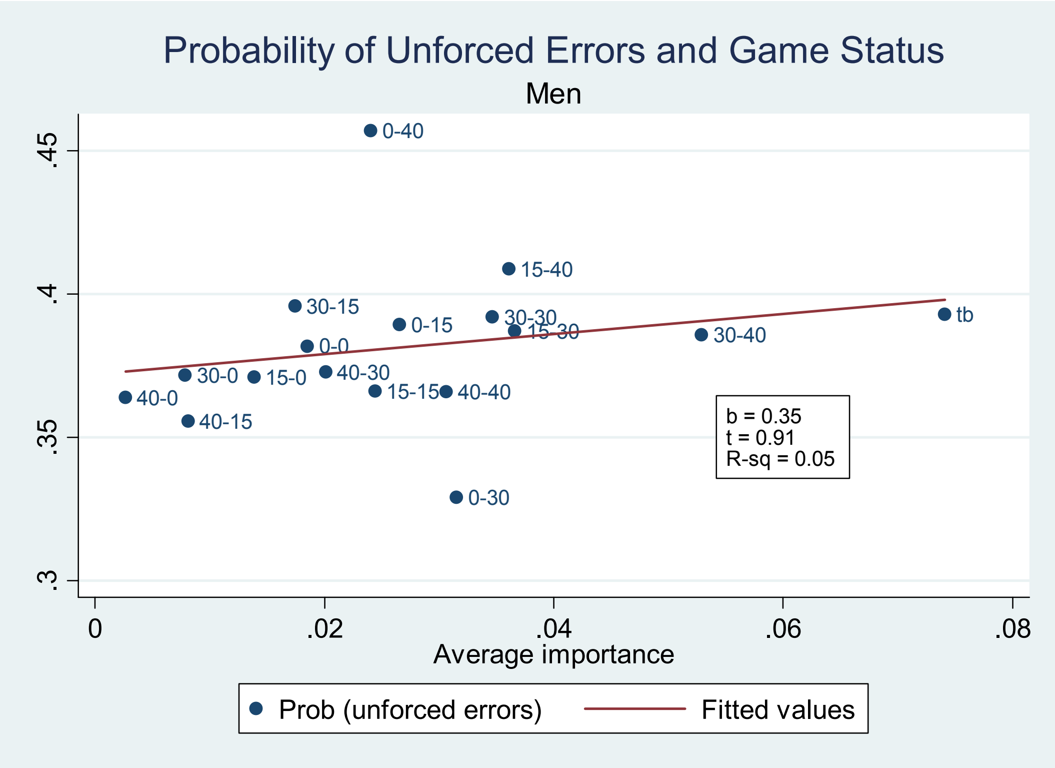


Figure 3a: Probability of Unforced Errors and Game Status – Men

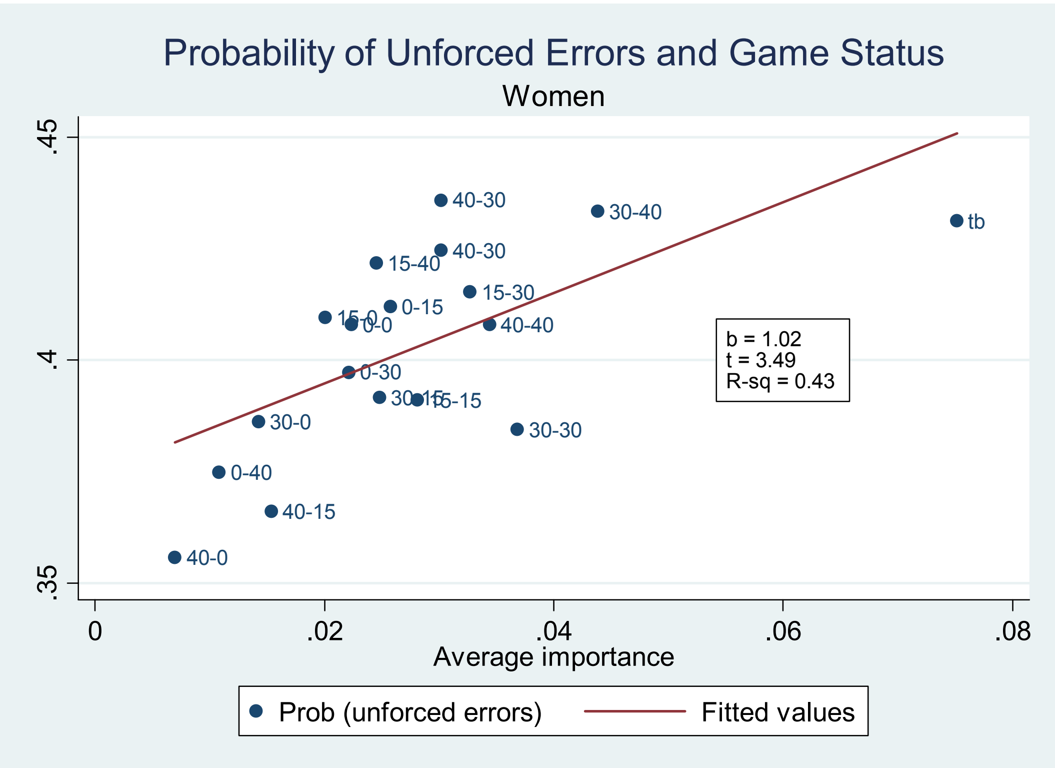


Figure 3b: Probability of Unforced Errors and Game Status – Women