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

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

May 2007

Abstract

This paper uses data from nine tennis Grand Slam tournaments played between 2005 and 2007 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 four of the nine 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 serve speed, on first serve percentages and on rally length suggest that women play a more conservative and less aggressive strategy as points become more important. I present a simple game-theoretic model that shows that a less aggressive strategy may be a player's best response to an increase in the intrinsic probability of making unforced errors.

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

* I thank Dan Ariely, Iván Fernández-Val, Esteban Klor, Motty Perry, 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 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

[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/). All these websites were last visited on May 14, 2007.

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 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.

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 in the nine Grand Slam tournaments played between January 2005 and January 2007. 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, Weinberg et al., 1981).

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 four of the nine 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, and the 2007 Australian Open, allows me to create a more 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. 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 36 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. Other data shows that on important points women hit slower first and second serves; have a higher first serve percentage, which, however, does not translate into a higher percentage of points won by the server; and play longer rallies. These results suggest that women play a more conservative and less aggressive strategy as points become more important. Since by playing a safe strategy they are also less likely to hit outright winners, the percentage of points ending in unforced errors rises. I also examine whether the differential reaction in performance in response to increases in the stakes can be explained by physical differences between men and women, but the evidence is inconclusive. A simple game-theoretic model shows that a shift from an aggressive to a less aggressive strategy may arise as an optimal response to a decrease in the intrinsic probability of making unforced errors. Simply put, if players know that on certain points they are more likely to make unforced errors, they will revert to a safer playing strategy.

While these empirical findings are certainly intriguing, one should be cautious before drawing from them broad generalizations about the operation of 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 make us even more aware of the dangers of extrapolation. Nevertheless, the fact that we find such a robust gender difference in performance under pressure even in the extreme right tail of the talent distribution is without a doubt a novel finding that should stimulate further investigation.

In addition to the recent experimental literature on performance in competitive environments, the paper is also linked to the vast literature in social psychology on 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), 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). Recently, Ariely et al. (2005) 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.

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 key variable in our analysis, the importance variable. Section 5 presents the basic results of the point-by-point analysis on the link between unforced errors and importance, while Section 6 looks at the link between importance and the aggressive play. Section 7 investigates whether the gender differences uncovered in the previous analysis are really attributable to gender *per se*, or whether they rather reflect differences in physical characteristics between men women. Section 8 presents a simple game-theoretic model of tennis points, including a simulation that illustrates how even small differences in performance at crucial stages of the game can have quite dramatic effects on the overall probability of winning the game. Section 9 discusses the results and 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. 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.[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 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 nine Grand Slam tournaments played between 2005 and 2007: all the tournaments played in 2005 and 2006, and the 2007 Australian Open, played in January 2007. The four Grand Slam tournaments (the Australian Open, the French Open, Wimbledon,

[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*, last viewed on April 26, 2007.

[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.

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 2006 and February 2007 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 seven of the nine 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 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 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 aggressiveness 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. On the other hand, a less aggressive strategy may also result in *more* points ending in unforced errors, if players refrain from going for winners, so that the point inevitably will end in an error by one of the two players.

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 3,727 sets in 1,023 matches for men, and 2,343 sets in 1,019 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. Predictably, men win a higher fraction of their service points, and for both men and women the percentage of points won by the server increases with the speed of the surface.

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:

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

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.

[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).

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 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 performance *within* a match, i.e., whether the quality of the 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

[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.

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 panel presents the results using the final set dummy as the measure of the magnitude of the stakes, while the bottom panel uses 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 one controls for match fixed effects. The percentage of unforced errors rises by 1.32 – 1.55 percent for men, and by 2.58 – 2.72 percent for women. The percentage of winners drops for both genders: by 1.2-1.4 percent (men), and by about 1.6 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 1 percent level, while the coefficients for men are statistically significant at the 5 percent level 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.

The bottom two rows of each panel present results for the effect of the magnitude of the stakes on the percentage of valid first serves and on the percentage of points won by the server. The percentage of valid first serves does not seem to vary in the final set of the match for either men or women. There is some evidence that women hit a higher percentage of first serves when the stakes, as measured by the importance variable, are high, but this result is not robust to the inclusion of match fixed effects. Interestingly, though, the higher percentage of first serves hit by women does not translate into a higher percentage of points won by the server. This suggests that maybe women adopt a less aggressive strategy at important stages of the match. We will see further evidence for this conjecture in the point-by-point analysis.

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. Table 3 presents 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 competitive 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 four tournaments in the sample, the 2006 French Open, Wimbledon, and US Open, and the 2007 Australian 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.[7] 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 238 matches, 127 for men and 111 for women. Altogether, I have data for nearly 42,000 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:

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

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

[7] See *http://www.rolandgarros.com/en_FR/pointtracker/index.html* for a description of the Point Tracker technology at the 2007 French Open.

probabilities are calculated using the full data from the remaining five tournaments, for which detailed point-by-point data was not available. I ran regressions of the proportion of points won by the server (the receiver) in each match on the rating of the two players and on tournament dummies, separately by gender. I then used the estimated coefficients from these regressions to estimate the predicted probability of each player winning a point for every match in the 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 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.

Table 4 presents summary statistics for the importance variable, separately by tournament and gender. The mean of the importance variable is 0.0235 for men, and 0.0297 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 = 0.086) 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 0.0437, about twice as large as when the score is 0-0 (0.0216), and more than 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 0.0044, compared to 0.0520 when the score is 30-40.

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.

5. Point-by-point analysis: results

Figure 3 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 30 and 31 percent, regardless of the importance of the point. On the other hand, the typology of points in women's matches is strongly affected by the importance variable. 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 three 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_{mt} as the outcome of point t in match m. Y_{mt} can take on three possible values: 1 – forced error, 2 – unforced error, 3 – winner. I estimate the following multinomial logit model:

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

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 a model in which the importance variable is included as a piecewise constant function, 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 6. The top panel presents the coefficients on the importance variables in the two models (linear 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 piecewise constant specification, however, reveals that there are some significant non-linearities in the effect of importance that are not captured in the linear specification. For men, all the coefficients on the importance quartile dummies are small and statistically insignificant in the no-fixed effects specification, and negative in the specification with match fixed effects. For women, on the other hand, the coefficients on the second and third importance quartile dummies are always positive and of moderate size, albeit not statistically different from zero. Strikingly, though, the coefficient on the fourth importance quartile dummy is large and positive, indicating that women are significantly more likely to hit unforced errors at the very crucial stages of the match. As a result, the gender difference in the propensity to hit unforced errors in the top quartile of the importance distribution is also large and statistically different from zero. Women also appear to hit relatively more winners in the highest importance quartile, but this finding is not robust to the inclusion of match fixed effects. It is also somewhat difficult to infer much about performance from this coefficient, since the probability of hitting a winner depends in part on the shot that preceded it.

The size of the coefficient implies that the odds of women making unforced errors (relative to forced errors) rise by 20-25 log points when moving from the bottom to the top quartile of the importance distribution. In terms of predicted probabilities, this implies that the probability of making unforced errors in the top quartile of the importance distribution rises by 3.8 – 4.3 percentage points.[8] 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.

Robustness checks

In Table 7 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

[8] The predicted probabilities are calculated for the 65th point played in the first set of a fourth round match between players ranked 15 and 16, on the main court of the French Open.

Table 7 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 6: the odds of women making unforced errors rises by 15-18 log points when moving from the bottom to the top quartile of the importance distribution. The coefficient is statistically significant at the 1 percent level in the specification without fixed effects, and marginally significant at the 10 percent level in the specification with fixed effects. The gender difference is marginally insignificant in both specifications.

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 7: the multinomial logit model (analogous to that of Table 6) 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, marginally significant in the logit model 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.

The evidence from Figure 3, as well as the comparison between the multinomial logit and simple coefficients in the previous tables, suggests that the higher incidence of unforced errors at important points in the women's game is accompanied by a decrease in the incidence of forced errors. Since "... most missed returns of first serves are considered to be forced errors – forced by the pace and placement of the opponent's serve,"[9] it is possible that the decrease in forced errors and the increase in unforced errors is simply a result of the fact that on important points women hit slower first serves (we will see that this is indeed the case in the next section), and are thus less likely to induce a forced error. Therefore, one would like to examine whether the increase in unforced errors occurs also if one focuses exclusively on long rallies, where the speed of the first serve is not a factor. In Table 8, I look at the effect of importance on the propensity to make unforced errors in the sample of points that include three or more strokes (i.e., points that do not end with a service or return winner or error). The top panel presents the coefficients in the "unforced errors" equation in a multinomial logit model, and the bottom panel presents the coefficients in a logit model, in which the dependent variable is a binary indicator for whether the point ends in an unforced error. The effect of importance on women's propensity to make unforced errors is even larger in this restricted sample of long rallies, and the difference between men and women is always statistically significant. Hence, it appears that the increase in the percentage of unforced errors at important points is not just a mechanical artifact of the fact that women hit slower serves.

As a final robustness check, I present in Figure 4 an alternative way of looking at the relationship between importance and the propensity to make unforced errors for the two genders, which focuses exclusively on variations in performance at different stages within the *game*, and illustrates more transparently the relationship between performance and the importance variable. Specifically, I estimated a multinomial logit model as in Table 6, 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

[9] http://www.tennis.com/yourgame/asktheeditors/asktheeditors.aspx?id=1432.

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. [10]

6. Importance and aggressiveness

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 men and women adopt different levels of aggressiveness as points become more important. One must be careful, because the effect of a more aggressive strategy (i.e., trying to produce more outright winners by hitting more powerful shots or trying to hit the corners of the court) on the percentage of points that end in unforced errors is ambiguous. On one hand, more aggressive shots are riskier, in the sense that they have a high probability of ending in unforced errors as well as winners. On the other hand, if players just hit the ball softly back and forth without trying to hit winners, almost by definition the point will end eventually with an unforced error. In this section, I use data on the speed and the accuracy of serves, and on the length of rallies, to assess whether differences in aggressiveness can explain the gender difference in the propensity to make errors at the crucial stages of the match.

The results are presented in Table 9. The table is divided into five panels, each one showing the results of linear regressions with different dependent variables. All the regressions control for the same set of explanatory variables as in Table 6.

Panel A of the table looks at the effect of the importance quartile on the speed of the first serve.[11] We observe substantial gender differences in serve speed as the stakes become

[10] In all of the above, I have assumed all the time that the importance of a point is the same for 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. I have conducted an analysis in which I decompose the importance of a point into the *anticipation of winning* and the *fear of losing*. The results are broadly consistent with those of Table 6, and 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. The results are available upon request.

[11] Serve speed is available only for about 77 percent of all points available in the point by point data. This explains the smaller combined number of observations in the serve speed regressions.

higher. Men hit 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.[12] Panel B of the table shows that 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.

Panels C and D provide additional evidence that women adopt a less aggressive strategy as points become more important. Panel C shows the effect of importance on the probability of hitting a valid first serve, using a linear probability model. For men, this probability is roughly constant regardless of the importance of the point. By contrast, women hit a higher percentage of first serves in the top two quartiles of the importance distribution: the probability of hitting a valid first serve rises by 3.5-4.5 percentage points in the third quartile, and by 1.4-2.5 percentage points in the fourth quartile (with the coefficient in the fixed-effects specification being statistically insignificant). Interestingly, in comparing the third and fourth quartiles of the importance distribution, we see that women's first serves are both less powerful and less accurate in the most important points. This result is very similar to that of Butler and Baumeister (1998), who found that performance anxiety elicited cautious, protective strategies that were associated with poor performance, i.e., decreases in speed without an associated improvement in accuracy.

A higher percentage of first serves would suggest that women's performance actually increases as points become more important. However, hitting a valid first serve is no guarantee of winning the point. In panel D of the table, I look at whether the server is more likely to win the point as importance increases. Surprisingly, the higher percentage of first

[12] 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, if one assumes that serve speed is negatively correlated with accuracy, it is straightforward to show that the likely direction of the 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.

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). The combined evidence of panels A-D indicates that women hit “softer” first serves, which are more likely to be in, but do not convey a significant advantage to the server. In other words, women adopt a more conservative playing strategy at the crucial stages of the match.

The last panel of the table looks at the effect of importance on the number of strokes in the rally. If women adopt a less aggressive strategy, as it appears from the analysis of the speed of the serve, then they should also be involved in longer rallies. The results show quite unambiguously that women play longer rallies as the importance of the point increases. Points in the top quartile of the importance distribution are between two thirds of a stroke and a stroke longer than points in the bottom quartile. Coupled with the slower and less effective first serves, this finding supports the hypothesis that women play more conservative strategies as points become more important. This can explain the higher incidence of points that end in unforced errors. If both players are just trying to put the ball in play without attempting to force the point, inevitably the point will end with one of the two players making a mistake.

The striking finding of Panel E, though, is that importance has an even stronger positive effect on the length of men’s rallies. This finding is more difficult to interpret, given that the evidence on serves for men is mixed and does not allow to infer a clear-cut relationship between the importance of the point and aggressiveness. One possible explanation is that men exert more effort on important points and try to chase down all balls, thus sustaining the rally. This would be consistent with the experimental evidence that men exert more effort in high stakes situations (Gneezy, Niederle and Rustichini, 2003). Of course, this explanation could also hold for women, but it is difficult to make a firmer statement without information on the actual strength, depth, and angle of each stroke.

7. Physical Differences and Performance

The previous analysis has shown that women adopt a less aggressive strategy and make more unforced errors on important points, while we find no such evidence for men. It is natural to ask whether these differences are really attributable to gender *per se*. After all, the differences in power, speed and other physical attributes between male and female professional tennis players are so large that it could well be that it is these physical differences that generate the different reactions to high stakes situations. For example, it is possible that on crucial points all players tend to adopt the playing strategy with which they are most comfortable: for women, who are less likely to overwhelm their opponents by sheer power, this means moving to a less aggressive strategy, which eventually leads to more unforced errors.

I address this hypothesis by conducting a series of *within-gender* analyses, where I classify players as high-power and low-power, based on their physical characteristics. I use two different classifications of high and low power players: the first is based on the average first serve speed, and the second is based on height. For both classifications, players are classified as high-power if they are above the gender and tournament specific median of the relevant variable, and low-power otherwise. I then ask, separately for men and women, whether there are any differences between the high-power players and the low-power players in the propensity to make unforced errors and in aggressiveness (as measured by first serve speed) as points become more important. The results are presented in Table 10.

The top panel presents the coefficients of the "unforced error" equation in the multinomial logit model. In this model, the low-power sample includes all points played between two players who are defined as "low-power", while the high-power sample includes all the other points (either points played by two high-power players, or points between a high-power and a low-power player). To preserve space, I present only the coefficient on the 4th importance quartile. The first row shows that there is no difference in the percentage of unforced errors between the top and the bottom quartile of importance, both for points played between low-power male players (in this case, players with low average first serve speed), and for all other points. The second row suggests that the proportion of unforced errors increases in the fourth quartile of the importance variable in matches between two short players, while it actually decreases in all other matches. This finding is consistent with the hypothesis that low-power players tend to make more unforced errors at crucial points, even though the coefficients are not always statistically significant. The third and fourth rows present the results for women: for both definitions, we find that the coefficients are larger in the low-power sample, but unforced errors increase also among the high power sample, and the difference between the two is mostly insignificant.

The bottom panel studies whether high-power and low-power players modify the speed of their serves as points become more important. Men with high average first serve speed significantly raise the speed of their first serve on important points, but there is no such effect for players with low average first serve speed. On the other hand, there is no difference between short and tall players in first serve speed on important points. For women, there is some evidence that it is mostly the low power women that reduce their first serve speed on important points, but the results are sensitive to the exact definition high and low power players and to the econometric specification.

Summing up, even though some of the results point in the direction that all low-power players adopt a less aggressive strategy and make more unforced errors on important points irrespective of gender, the evidence is far from conclusive, and it does not contradict the basic finding that there are substantial differences in the way men and women approach important points in the match.[13]

[13] An alternative approach to studying the effects of personal characteristics on performance under pressure involves estimating a full set of individual-specific propensities to make unforced errors on important points, and then regressing these estimated coefficients on the player's personal characteristics. These regressions yielded mostly insignificant results, apart from the somewhat unsurprising finding that higher ability players are less likely to make mistakes under pressure.

8. A Simple Model of Tennis Points

The empirical results paint a fairly clear and consistent picture. At crucial stages of the match, female players adopt a less aggressive strategy, and, as a consequence, it is more likely that points will end in unforced errors by either one of the players. Is it accurate, though, to depict this change in strategy as a deterioration in performance? We must recognize that tennis is a game that involves strategic interactions: it is possible that at different stages of the match players choose different equilibrium strategies, which are not necessarily suboptimal. In the following, I present a simple game-theoretic model of a tennis point, which illustrates that a shift from a relatively aggressive to a relatively "soft" strategy can only occur if there is a change in the intrinsic probabilities of hitting winners or committing unforced errors.

The model

I model here 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), 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.

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 us to concentrate on the salient aspects of the game that we are interested in.

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. For example,

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

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:

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

and the action chosen by player B satisfies:

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

To further simplify the model, I make the additional assumption that the two players have identical abilities, hence $w_{1A} = w_{1B} = w_1$; $w_{0A} = w_{0B} = w_0$; $u_{1A} = u_{1B} = u_1$; $u_{0A} = u_{0B} = u_0$.[14] One can now state Proposition 1.

Proposition 1: 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. 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:

[14] 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.

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

Proof: See Appendix.

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

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

which can also be rewritten as:

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

Proof: See Appendix.

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

Proof: See Appendix.

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: See Appendix.

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.*

The implication of Proposition 1 is clear: if different types of equilibria arise in 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. Note that the inequalities reflect 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. The inequalities 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.

Linking this back to the empirical analysis, this means that playing a less aggressive style of play at more important points of the match may be simply the optimal equilibrium response to an increased probability of committing unforced errors (or a decreased probability of hitting winners) when playing aggressively on those points. To illustrate this point, I present two numerical examples.

Example 1: Assume that the intrinsic probabilities are: $w_0 = 0.096;\ w_1 = 0.12; u_0 = 0.064; u_1 = 0.07.$[15] Since $(w_1 - w_0) + u_1(p_1 - p_0) + p_1(w_0 p_1 - w_1 p_0) > 0$, the unique subgame perfect equilibrium in this game is (Aggressive, Aggressive), and the equilibrium payoffs are $V^A(1,1) = 0.5138$ and $V^B(1,1) = 0.4862$.[16]

Example 2: Assume that the intrinsic probabilities are: $w_0 = 0.096; w_1 = 0.10; u_0 = 0.064; u_1 = 0.07.$ Since $(w_1 - w_0) + u_0(p_1 - p_0) + p_0(w_0 p_1 - w_1 p_0) < 0$, the unique subgame perfect equilibrium in this game is (Soft, Soft), and the equilibrium payoffs are $V^A(0,0) = 0.5087$ and $V^B(0,0) = 0.4923$.

[15] These probabilities, as well as those in the next examples, are chosen so that the percentage of points that end in unforced errors and the length of the rally roughly match those observed in the data.

[16] $V^B(1,1)$ is the payoff to player B when player A starts the rally. Note that $V^B(1,1) = u_1 + p_1 V^A(1,1)$, i.e., it is identical to the probability that player A makes an unforced error plus the probability that the ball reaches player B, who then follows the same optimal strategy as player A when it was his turn to start the rally.

It is interesting to evaluate what would happen in this game if, for example, player A could reduce the probability of making unforced errors when playing aggressively from 0.6 to 0.4, while player B continues to have the same probabilities as in Example 2.

Example 3: Assume that the intrinsic probabilities are: $w_{0A} = 0.096; w_{1A} = 0.12; u_{0A} = 0.064; u_{1A} = 0.07;$ and $w_{0B} = 0.096; w_{1B} = 0.10; u_{0B} = 0.064;$ $u_{1B} = 0.07$. In a game with heterogeneous players, the conditions for a unique (Aggressive, Aggressive) subgame perfect equilibrium are:

$$\textbf{C1':}\ (w_{1A} - w_{0A}) w_{1B} - (u_{1A} - u_{0A}) u_{1B} - (1 - w_{1B} - u_{1B})(w_{0A} u_{1A} - w_{1A} u_{0A}) > 0,$$

and

$$\textbf{C1'':}\ (w_{1B} - w_{0B}) w_{1A} - (u_{1B} - u_{0B}) u_{1A} - (1 - w_{1A} - u_{1A})(w_{0B} u_{1B} - w_{1B} u_{0B}) > 0.$$

These conditions are satisfied here, hence the unique subgame perfect equilibrium is (Aggressive, Aggressive), and the payoffs are $V^A(1,1) = 0.5392$ and $V^B(1,1) = 0.4608$. Relative to Example 2, player A increases the probability of winning the point from 0.5087 to 0.5392. Similarly, we can show that if player B (the second player to strike the ball) raises his or her game by increasing the probability of hitting winners when playing aggressive from 0.10 to 0.12, the probability of winning the point goes from 0.4923 to 0.5175. To assess how important these changes are, I perform a small simulation exercise, whose results are presented in Table 11. In the benchmark case, both players have intrinsic probabilities as in Example 1 on non-important points, and intrinsic probabilities as in Example 2 on important points. In the treatment case, the intrinsic probabilities of player A become those of Example 3, i.e., player A is able to avoid the drop in performance on important points. The increase in player A's probability of winning the entire match can be quite dramatic, going from 0.5 to 0.710 in three-set matches, and 0.787 in 5 set matches. In other words, a player that could control the tendency to make unforced errors at 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.

9. Conclusion

In this paper I have used data from nine Grand Slam tournaments played between January 2005 and January 2007 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 junctures 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. Some of this difference can be explained by gender differences in type of play as points become more important: the evidence on rally length and on the speed and accuracy of first serves strongly suggests that women tend to adopt a safer and less aggressive strategy on important points. 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 control their tendency to make unforced errors on important points can substantially increase their chances of winning.

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 and execution by the player. Can we extrapolate from our findings that in general women's judgment and execution 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. In fact, beyond the obvious dangers of taking any study out of its context, there are a number of specific features of tennis that should make us even more cautious with the interpretation of the results.

First, the sample of professional tennis players under examination is an extremely selected sample, and it is possible that the selection mechanism operates differently for men and women. For example, men with a high predisposition to excel in sports have a wide variety of disciplines from which to choose from, and in which it is possible to reap substantial financial rewards. The set of disciplines that are financially rewarding for women is much more limited. In addition, it may be more difficult for men to emerge through the youth system and enter the professional circuit, since the competition among boys is stiffer.[17] Given these two factors, it is not inconceivable that men who reach the top levels of professional tennis are an even more select group relative to women, especially in their ability to perform well in high pressure situations.

Second, performance in tennis draws primarily on motor skills, contrary to most of the tasks that are relevant for the very high skill workers for which large gender gaps still persist. It is possible that the mechanisms that induce decreased performance under pressure in tasks involving motor skills are different from mechanisms that are relevant for cognitive tasks. For example, Baumeister (1984) argues that decreased performance under pressure in tasks involving motor skills arises because subjects begin to consciously think of the actions necessary to perform the task, and shift control from automatic processes to less effective

[17] In conversations I have had with a number of tennis coaches, it has been suggested that young girls are substantially more likely to drop out of competitive tennis as a result of injuries, financial difficulties, a drop in motivation, or pursuit of other interests. Deaner (2006a and 2006b) finds that both among elite track runners and road runners, two to four times as many males as females run fast relative to sex-specific world-class standards.

controlled processes (this is known as the *explicit monitoring* hypothesis). Beilock and Carr, (2005), on the other hand, argue that performance anxiety may negatively affect working memory capacity available for skill execution even in math problems that do not involve much in the way of proceduralized routines. One must be aware of these different mechanisms when attempting to draw lessons from this paper's findings for the labor market in general, and for high skill workers in particular.

Third, even putting aside the differences between tasks involving motor skills and cognitive skills, the set of professions in which workers are required to make the type of accurate high-stakes split-second decisions that are under examination in this study is fairly limited: maybe floor traders, emergency room surgeons, combat military personnel, and few others. To what extent then can the results in this study be used to explain the under-representation of women among elite lawyers, corporate managers and academic researchers? These are clearly professions that involve many high-pressure situations, but it is probably pressure of a different type than that faced by tennis players.

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.

Summing up, this study has uncovered a striking empirical regularity on gender differences in performance under pressure among elite professional tennis players. Putting aside all the aforementioned caveats, and assuming that the empirical finding does indeed reflect a true gender difference in performance under pressure, can it explain the gender differences in representation at the highest rungs of the occupational ladder? There are at least two possibilities. The first, and more controversial one, is that this is a real gender difference in productivity that is relevant for many contexts other than tennis, and therefore profit-maximizing employers may refrain from hiring women for very top positions. Alternatively, it is possible that performance under pressure is only one aspect of productivity on the job, and maybe not the most important one.[18] Maybe what hinders the advancement of women is the nature of the internal promotion tournaments within the firm, which are more resembling of sports contests. Either because they shy away from competition, or they perform less well under pressure, or they act less aggressively, women find it harder to emerge on top of these tournaments: this is the reason for their under-representation in the top positions of the organization.

Of course, it is dangerous to extrapolate from a single study, based on a very select group of individuals who engage in an activity that is in many ways substantially different from those encountered in the business world or in academia. Yet, the findings are sufficiently interesting that they should stimulate further research and a deeper understanding of this matter.

[18] For example, Babcock and Lascehever (2003) argue that women can reach superior outcomes in certain types of negotiations because they tend to behave more cooperatively than men.

A1. Proof of Proposition 1

Proof of Lemma 1: 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:

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

Proof of Lemma 2: The pair of strategies (Soft, Soft) 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:

$$V^A(1,1) = w_1 + p_1 u_1 + p_1^2 V^A(1,1),$$
$$\Rightarrow V^A(1,1) = \frac{w_1 + p_1 u_1}{1 - p_1^2};$$

and:

$$V^A(0,1) = w_0 + p_0 u_1 + p_0 p_1 V^A(0,1),$$
$$= \frac{w_0 + p_0 u_1}{1 - p_0 p_1}.$$

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

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

Rearranging terms, we have:

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

Similarly, we have that

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

and:

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

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

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

Rearranging terms, we have:

(2b) $$(w_1 - w_0) + u_1(p_1 - p_0) + p_1(w_0p_1 - w_1p_0) > 0,$$

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

Proof of Lemma 3: 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 $u_1 = \dfrac{u_0^2 + w_0(w_1 - w_0) + (1 - u_0 - w_0)w_1u_0}{u_0 + (1 - u_0 - w_0)w_0}$ (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.■

Proof of Lemma 4: 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.■

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.

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.

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.

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." NBER Working Paper No. 10678, August 2004.

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.

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: Summary Statistics

	Men					Women				
	Total	Australian Open	French Open	Wimbledon	US Open	**Total**	Australian Open	French Open	Wimbledon	US Open
Number of matches	1023	377	249	244	153	1019	379	254	245	141
Number of sets	3,727	1378	900	876	573	2,343	868	586	570	319
Pct. unforced errors	32.63 (11.70)	33.29 (9.55)	41.96 (11.33)	22.22 (7.88)	32.30 (8.46)	40.72 (12.07)	43.96 (9.76)	47.45 (11.73)	30.11 (8.94)	38.48 (9.89)
Pct. winners	34.38 (7.95)	33.66 (8.19)	34.12 (7.50)	35.54 (7.81)	34.72 (8.03)	29.97 (7.94)	28.27 (7.84)	31.43 (7.80)	30.89 (7.29)	30.27 (8.76)
Pct. forced errors	33.00 (10.73)	33.05 (8.02)	23.92 (10.21)	42.24 (8.37)	32.99 (8.00)	29.31 (11.01)	27.77 (7.91)	21.11 (10.57)	38.99 (9.08)	31.25 (8.18)
Pct. first serve	60.60 (10.84)	59.24 (10.64)	61.59 (11.28)	62.25 (10.36)	59.76 (10.87)	61.63 (11.28)	60.21 (10.91)	61.94 (11.61)	63.09 (11.42)	62.31 (11.04)
Pct. points won by server	61.92 (12.66)	60.96 (12.81)	59.71 (12.73)	65.06 (11.32)	62.93 (13.09)	55.52 (13.25)	55.07 (13.33)	53.82 (12.91)	57.63 (12.83)	56.06 (13.92)
Average player rank	83.88	84.04	83.86	87.31	78.04	80.77	83.90	80.96	82.74	68.97
Average player rating	2.23	2.20	2.19	2.15	2.47	2.26	2.20	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: Stakes variable: final set dummy			
Dependent variable:						
Pct. unforced errors	1.3207	2.5779	1.2572	1.5490	2.7209	1.1719
	[1.86]	[3.31]	[1.19]	[2.26]	[3.56]	[1.14]
Pct. winners	-1.1900	-1.6321	-0.4420	-1.4027	-1.5787	-0.1760
	[-1.93]	[-2.67]	[-0.51]	[-2.33]	[-2.59]	[-0.21]
Pct. first serve	-0.2614	0.2296	0.4910	-0.0378	-0.4098	-0.3719
	[-0.44]	[0.36]	[0.57]	[-0.07]	[-0.65]	[-0.44]
Pct. won server	0.2485	-0.8737	-1.1222	0.0791	-0.9603	-1.0394
	[0.44]	[-1.43]	[-1.34]	[0.13]	[-1.52]	[-1.20]
			B: Stakes variable: importance of the set			
Dependent variable:						
Pct. unforced errors	1.6360	2.1385	0.5024	0.9229	2.6862	1.7633
	[1.81]	[1.82]	[0.34]	[1.03]	[2.29]	[1.19]
Pct. winners	-1.9559	-1.5029	0.4530	-2.3084	-1.8750	0.4334
	[-2.46]	[-1.54]	[0.36]	[-2.91]	[-1.98]	[0.35]
Pct. first serve	-0.3788	2.1098	2.4887	0.2813	0.3333	0.0520
	[-0.50]	[2.36]	[2.13]	[0.35]	[0.34]	[0.04]
Pct. won server	1.1632	-0.9007	-2.0639	0.6225	-1.0455	-1.6680
	[1.55]	[-1.04]	[-1.81]	[0.78]	[-1.05]	[-1.31]
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of observations	3,727	2,343	6,070	3,727	2,343	6,070

Note: Entries in the table represent the coefficient on the stakes variable, 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)
	Dependent variable: percent unforced errors							
Final Set	1.300 [1.83]	1.377 [1.96]	1.514 [2.21]	1.557 [2.31]	2.553 [3.28]	2.697 [3.43]	2.694 [3.52]	2.697 [3.52]
Total number of points played before start of set	-0.010 [-4.51]	-0.010 [-4.72]	-0.009 [-4.75]	-0.010 [-4.87]	-0.020 [-3.59]	-0.022 [-3.83]	-0.023 [-4.35]	-0.023 [-4.29]
Final Set × number of points (centered)	-	0.043 [1.97]	-	0.043 [2.20]	-	0.022 [0.83]	-	0.001 [0.02]
Match fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Number of observations	3,727	3,727	3,727	3,727	2,343	2,343	2,343	2,343

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	Australian Open 2007	All	French Open 2006	Wimbledon 2006	US Open 2006	Australian Open 2007	All	**ALL**
Mean	0.0220	0.0243	0.0255	0.0220	0.0235	0.0316	0.0236	0.0226	0.0387	0.0297	**0.0256**
Standard deviation	0.0299	0.0329	0.0373	0.0291	0.0327	0.0411	0.0343	0.0363	0.0486	0.0413	**0.0360**
25^{th} percentile	0.0023	0.0030	0.0028	0.0021	0.0026	0.0052	0.0013	0.0004	0.0071	0.0029	**0.0027**
50^{th} percentile	0.0116	0.0129	0.0133	0.0120	0.0125	0.0170	0.0092	0.0060	0.0230	0.0142	**0.0129**
75^{th} percentile	0.0305	0.0322	0.0338	0.0305	0.0318	0.0440	0.0314	0.0303	0.0532	0.0420	**0.0348**
Number of points	7,776	7,076	7,529	5,260	27,641	4,728	2,587	3,532	3,477	14,324	**41,965**

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	0.0111	0.0111	.	5-0	0.0014	40-0	0.0044	0.0028	0.0087
0-0	0.0192	0.0176	0.0212	0-5	0.0018	30-0	0.0102	0.0081	0.0154
1-0	0.0228	0.0192	0.0272	1-5	0.0052	40-15	0.0105	0.0081	0.0160
2-1	0.0331	0.0331	.	0-4	0.0065	15-0	0.0164	0.0140	0.0219
1-1	0.0508	0.0352	0.0727	2-5	0.0089	0-40	0.0191	0.0231	0.0145
2-2	0.0855	0.0855	.	5-1	0.0093	0-0	0.0204	0.0185	0.0242
				4-0	0.0095	30-15	0.0207	0.0176	0.0272
				5-2	0.0100	40-30	0.0247	0.0202	0.0332
				1-4	0.0104	15-15	0.0264	0.0241	0.0306
				0-3	0.0124	0-15	0.0267	0.0262	0.0275
				3-5	0.0144	0-30	0.0278	0.0307	0.0238
				3-0	0.0155	15-40	0.0323	0.0351	0.0287
				1-3	0.0156	deuce	0.0352	0.0323	0.0401
				2-4	0.0161	15-30	0.0352	0.0352	0.0352
				0-2	0.0177	30-40	0.0520	0.0531	0.0504
				4-1	0.0182	TB	0.0720	0.0699	0.0826
				1-0	0.0205				
				2-0	0.0206				
				0-1	0.0212				
				0-0	0.0216				
				3-1	0.0229				
				2-1	0.0229				
				1-2	0.0238				
				1-1	0.0242				
				4-2	0.0266				
				2-2	0.0267				
				4-3	0.0287				
				2-3	0.0293				
				3-2	0.0299				
				3-3	0.0312				
				5-3	0.0331				
				3-4	0.0351				
				5-4	0.0386				
				4-4	0.0388				
				5-6	0.0411				
				5-5	0.0437				
				4-5	0.0447				
				6-5	0.0548				
				TB	0.0720				
				6-6	0.2270				

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: 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	0.1464	1.2170	1.0706	-0.3165	0.4703	0.7868
	[0.22]	[1.54]	[1.04]	[-0.47]	[0.56]	[0.73]
B: Piecewise constant						
Importance	-0.0107	0.0694	0.0801	-0.0648	0.0916	0.1563
quartile 2	[-0.21]	[0.92]	[0.89]	[-1.11]	[1.14]	[1.58]
Importance	0.0065	0.0772	0.0707	-0.0753	0.0153	0.0907
quartile 3	[0.11]	[0.99]	[0.73]	[-1.17]	[0.15]	[0.74]
Importance	0.0226	0.2531	0.2306	-0.0750	0.1959	0.2710
quartile 4	[0.39]	[3.39]	[2.44]	[-1.02]	[1.85]	[2.10]
Winners						
A: Linear						
Importance	-0.8084	0.4130	1.2214	-1.0847	-0.5271	0.5576
	[-1.37]	[0.62]	[1.38]	[-1.95]	[-0.66]	[0.57]
B: Piecewise constant						
Importance	-0.0471	0.0024	0.0495	-0.0696	0.0586	0.1282
quartile 2	[-0.93]	[0.04]	[0.58]	[-1.22]	[0.77]	[1.35]
Importance	-0.0612	0.0281	0.0893	-0.0979	0.0003	0.0982
quartile 3	[-1.14]	[0.36]	[0.94]	[-1.60]	[0.00]	[0.81]
Importance	-0.0477	0.1348	0.1825	-0.0959	0.0795	0.1754
quartile 4	[-0.97]	[1.82]	[2.05]	[-1.58]	[0.76]	[1.45]
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of observations	27,595	14,305	41,900	27,595	14,305	41,900

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, US Open 2006, and Australian Open 2007 tournament dummies. The base category is "forced errors." Robust z-statistics (adjusted for clustering at the match level) in parentheses.

Table 7: 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 quartile 2	0.0131 [0.30]	0.0671 [1.05]	0.0540 [0.70]	-0.0298 [-0.54]	0.0605 [0.87]	0.0903 [1.02]
Importance quartile 3	0.0374 [0.73]	0.0601 [0.85]	0.0226 [0.26]	-0.0263 [-0.42]	0.0141 [0.16]	0.0405 [0.38]
Importance quartile 4	0.0465 [0.85]	0.1784 [2.75]	0.1319 [1.55]	-0.0271 [-0.38]	0.1528 [1.65]	0.1799 [1.54]
B: Importance not a function of players' ability rating: Multinomial logit coefficients on unforced errors (base category: forced errors).						
Importance quartile 2	0.0281 [0.66]	-0.0384 [-0.52]	-0.0665 [-0.78]	0.0236 [0.56]	-0.0036 [-0.05]	-0.0272 [-0.31]
Importance quartile 3	-0.0242 [-0.52]	0.0376 [0.51]	0.0618 [0.71]	-0.0163 [-0.36]	0.0516 [0.66]	0.0680 [0.75]
Importance quartile 4	-0.0194 [-0.36]	0.0693 [0.89]	0.0888 [0.94]	-0.0021 [-0.04]	0.0776 [0.94]	0.0797 [0.81]
C: Importance not a function of players' ability rating: Logit coefficients on binary dependent variable: unforced errors.						
Importance quartile 2	0.0565 [1.50]	0.0015 [0.02]	-0.0549 [-0.74]	0.0514 [1.35]	0.0249 [0.37]	-0.0265 [-0.34]
Importance quartile 3	-0.0085 [-0.21]	0.0905 [1.45]	0.0990 [1.33]	-0.0003 [-0.01]	0.0907 [1.36]	0.0910 [1.17]
Importance quartile 4	0.0420 [0.83]	0.1263 [1.95]	0.0842 [1.03]	0.0623 [1.25]	0.1275 [1.82]	0.0652 [0.76]
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of observations (points)	27,595	14,305	41,900	27,595	14,305	41,900

Note: Entries in the table are the coefficients in either a logit model (panel A) or a multinomial logit model (panels B and C) 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, US Open 2006, and Australian Open 2007 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 8: The Effect of Importance on Unforced Errors in Long Rallies

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
A. Unforced Errors (multinomial logit coefficients)						
Importance quartile 2	-0.0231 [-0.39]	0.1140 [1.32]	0.1371 [1.32]	-0.0694 [-1.00]	0.1012 [1.07]	0.1705 [1.46]
Importance quartile 3	-0.0174 [-0.26]	0.1316 [1.46]	0.1490 [1.34]	-0.1073 [-1.43]	0.0868 [0.70]	0.1941 1.34]
Importance quartile 4	0.0345 [0.53]	0.3356 [3.81]	0.3011 [2.75]	-0.0860 [-1.06]	0.2960 [2.34]	0.3820 [2.55]
B. Unforced Errors (logit coefficients)						
Importance quartile 2	0.0179 [0.36]	0.1003 [1.27]	0.0824 [0.88]	-0.0300 [-0.47]	0.0716 [0.82]	0.1017 [0.94]
Importance quartile 3	0.0112 [0.20]	0.0956 [1.05]	0.0844 [0.79]	-0.0709 [-1.00]	0.0548 [0.49]	0.1257 [0.95]
Importance quartile 4	0.0421 [0.06]	0.2250 [2.65]	0.1829 [1.75]	-0.0703 [-0.90]	0.2229 [1.96]	0.2931 [2.13]
Match fixed effects	No	No	No	Yes	Yes	Yes
Number of observations	17,788	9,474	27,262	17,788	9,474	27,262

Note: Entries in the table are the coefficients on the importance variables in the "unforced errors" equation in a multinomial logit model (panel A), and the coefficients in a logit model for the binary outcome "unforced errors" (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, US Open 2006, and Australian Open 2007 tournament dummies. The base category in the multinomial logit model is "forced errors." Robust z-statistics (adjusted for clustering at the match level) in parentheses.

Table 9: The Effect of Importance on Other Outcomes

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
A: Dependent variable: first serve speed (mph)						
Importance quartile 2	1.0309 [1.65]	-0.8003 [-1.54]	-1.8313 [-2.26]	1.2140 [3.59]	-0.7451 [-1.85]	-1.9591 [-3.74]
Importance quartile 3	0.4494 [0.61]	-1.7806 [-2.70]	-2.2300 [-2.26]	1.489 [3.64]	-1.2599 [-2.76]	-2.7490 [-4.50]
Importance quartile 4	0.9511 [1.17]	-2.7505 [-3.39]	-3.7016 [-3.23]	1.9509 [4.43]	-1.4302 [-2.65]	-3.3811 [-4.86]
Number of observations	14,722	7,631	22,353	14,722	7,631	22,353
B: Dependent variable: second serve speed (mph)						
Importance quartile 2	0.1218 [0.18]	-0.8156 [-1.07]	-0.9374 [-0.92]	-0.5825 [-1.51]	-1.4466 [-2.39]	-0.8642 [-1.21]
Importance quartile 3	-0.2527 [-0.35]	-2.0921 [-2.07]	-1.8394 [-1.48]	-1.0176 [-2.13]	-2.7489 [-3.97]	-1.7313 [-2.07]
Importance quartile 4	-0.7409 [-0.95]	-1.9170 [-1.90]	-1.1761 [-0.93]	-1.4684 [-3.01]	-3.3942 [-4.49]	-1.9258 [-2.15]
Number of observations	6,763	3,013	9,776	6,763	3,013	9,776
C: Dependent variable: first serve in						
Importance quartile 2	-0.0074 [-0.73]	0.0067 [0.54]	0.0141 [0.88]	0.0091 [0.78]	-0.0004 [-0.03]	-0.0095 [-0.52]
Importance quartile 3	0.0004 [0.04]	0.0450 [3.16]	0.0446 [2.49]	0.0115 [0.84]	0.0346 [1.96]	0.0231 [1.04]
Importance quartile 4	-0.0050 [-0.42]	0.0252 [2.05]	0.0302 [1.76]	0.0039 [0.28]	0.0137 [0.78]	0.0098 [0.44]
Number of observations	27,595	14,305	41,900	27,595	14,305	41,900
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, US Open 2006, and Australian Open 2007 tournament dummies. Robust t-statistics (adjusted for clustering at the match level) in parentheses.

Table 9: The Effect of Importance on Other Outcomes (cont'd)

	Individual Controls			Individual controls and match fixed effects		
	Men	Women	Difference	Men	Women	Difference
D: Dependent variable: server wins point						
Importance quartile 2	-0.0113 [-1.31]	-0.0159 [-1.27]	-0.0046 [-0.31]	-0.0214 [-1.93]	-0.0219 [-1.35]	-0.0005 [-0.03]
Importance quartile 3	-0.0137 [-1.51]	0.0010 [0.07]	0.0147 [0.91]	-0.0311 [-2.38]	-0.0146 [-0.83]	0.0165 [0.76]
Importance quartile 4	-0.123 [-1.06]	-0.0061 [-0.49]	0.0062 [0.37]	-0.0346 [-2.31]	-0.0263 [-1.42]	0.0083 [0.35]
Number of observations	27,595	14,305	41,900	27,595	14,305	41,900
E: Dependent variable: number of strokes per rally						
Importance quartile 2	0.1367 [1.22]	0.2438 [1.86]	0.1071 [0.62]	0.4362 [4.56]	0.4356 [3.55]	-0.0007 [-0.00]
Importance quartile 3	0.4371 [3.34]	0.1586 [1.15]	-0.2785 [-1.47]	0.9188 [8.08]	0.3847 [2.63]	-0.5341 [-2.89]
Importance quartile 4	0.8695 [5.55]	0.6640 [4.45]	-0.2055 [-0.95]	1.4589 [11.11]	0.9872 [6.32]	-0.4717 [-2.32]
Number of observations	24,355	12,186	36,541	24,355	12,186	36,541
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, US Open 2006, and Australian Open 2007 tournament dummies. Robust t-statistics (adjusted for clustering at the match level) in parentheses.

Table 10: Players' Physical Characteristics and Performance on Important Points

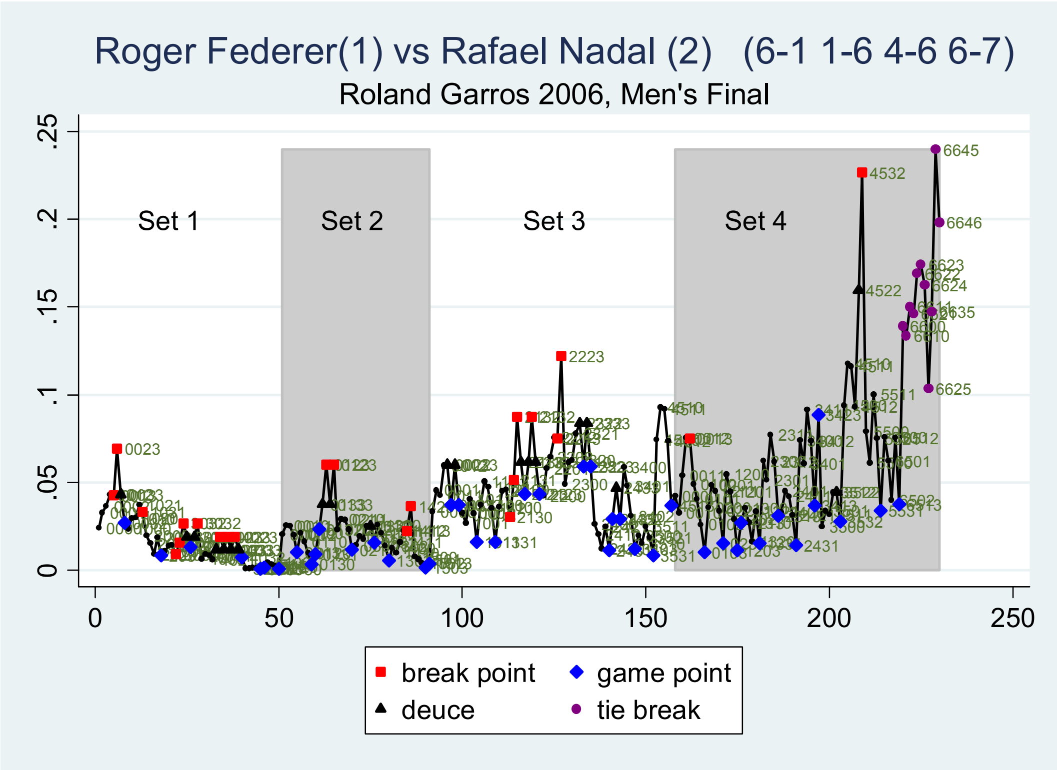
	Individual controls			Individual controls and match fixed effects		
	High Power	Low Power	Difference	High Power	Low Power	Difference
A: Unforced Errors: Multinomial logit coefficients on 4th quartile of the importance variable						
Men						
Power variable: average 1st serve speed	-0.0099 [-0.15]	0.1395 [0.95]	0.1494 [0.94]	-0.0931 [-1.02]	0.0083 [0.06]	0.1014 [0.59]
Power variable: height	-0.0911 [-1.27]	0.1957 [2.26]	0.2868 [2.56]	-0.1984 [-1.95]	0.0800 [0.76]	0.2784 [1.91]
Women						
Power variable: average 1st serve speed	0.1144 [1.33]	0.5481 [2.95]	0.4337 [2.14]	0.1899 [1.56]	0.2912 [1.20]	0.1014 [0.38]
Power variable: height	0.2191 [2.47]	0.3617 [2.37]	0.1426 [0.81]	0.1839 [1.52]	0.3712 [1.85]	0.1873 [0.81]
B: First serve speed: linear regression coefficients on 4th quartile of the importance variable						
Men						
Power variable: average 1st serve speed	2.8295 [4.25]	0.7111 [0.86]	-2.1184 [-1.99]	2.8071 [5.97]	0.4292 [0.62]	-2.3779 [-3.01]
Power variable: height	0.3587 [0.35]	0.4927 [0.60]	0.1340 [0.11]	1.9415 [2.99]	1.6317 [3.31]	-0.2424 [-0.33]
Women						
Power variable: average 1st serve speed	0.0362 [0.06]	-2.6883 [-2.79]	-2.7245 [-2.75]	0.1478 [0.26]	-1.0847 [-1.46]	-1.2324 [-1.34]
Power variable: height	-0.6706 [-0.73]	-1.9482 [-2.10]	-1.2777 [-1.05]	-0.6371 [-0.68]	-0.2675 [-0.51]	0.3696 [0.34]

Note: Entries in the table are the coefficients on the fourth quartile of the importance variable in the "unforced errors" equation in a multinomial logit model (panel A), and in a linear regression of first serve speed (panel B). In panel A, the "low power" sample includes all matches in which *both* players are below the median in terms of the power variable, and the "high power" sample includes all remaining matches. In panel B, the "low power" sample includes all points in which the server is below the median in terms of the power variable. 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, US Open 2006, and Australian Open 2007 tournament dummies. Robust t-statistics (adjusted for clustering at the match level) in parentheses.

Table 11: The effect of Raising One's Game on Important Points

	Importance < threshold				**Importance > threshold**					
	Player A's intrinsic probabilities	Player B's intrinsic probabilities	Equilibrium strategies	P(A wins point) [serving, receiving]	Player A's intrinsic probabilities	Player B's intrinsic probabilities	Equilibrium strategies	P(A wins point) [serving, receiving]	**P(A wins match), best of 3 sets**	**P(A wins match), best of 5 sets**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	**(9)**	**(10)**
Benchmark	$w_{0A} = 0.096$ $u_{0A} = 0.064$ $w_{1A} = 0.12$ $u_{1A} = 0.07$	$w_{0B} = 0.096$ $u_{0B} = 0.064$ $w_{1B} = 0.12$ $u_{1B} = 0.07$	(Aggressive, Aggressive)	0.5138, 0.4862	$w_{0A} = 0.096$ $u_{0A} = 0.064$ $w_{1A} = 0.10$ $u_{1A} = 0.07$	$w_{0B} = 0.096$ $u_{0B} = 0.064$ $w_{1B} = 0.10$ $u_{1B} = 0.07$	(Soft, Soft)	0.5087, 0.4923	**0.5**	**0.5**
Treatment	$w_{0A} = 0.096$ $u_{0A} = 0.064$ $w_{1A} = 0.12$ $u_{1A} = 0.07$	$w_{0B} = 0.096$ $u_{0B} = 0.064$ $w_{1B} = 0.12$ $u_{1B} = 0.07$	(Aggressive, Aggressive)	0.5138, 0.4862	$w_{0A} = 0.096$ $u_{0A} = 0.064$ $w_{1A} = 0.12$ $u_{1A} = 0.07$	$w_{0B} = 0.096$ $u_{0B} = 0.064$ $w_{1B} = 0.10$ $u_{1B} = 0.07$	(Aggressive, Aggressive)	0.5392, 0.5175	**0.710**	**0.787**

Note: The table presents the equilibrium strategies and payoffs for a single point according to the model presented in Section 8. On points with importance variable below the threshold, the intrinsic probabilities are presented in columns (1) and (2). On points with importance variable above the threshold, the intrinsic probabilities are presented in columns (5) and (6). The importance of each point is calculated assuming that players have identical abilities, and the threshold is the 75th percentile of the importance variable in the sample (from Table 4): 0.0318 for five-set matches, and 0.0420 for three set matches. Columns (9) and (10) present the implied probabilities of winning the entire match, given the intrinsic probabilities, and the equilibrium strategies and payoffs.



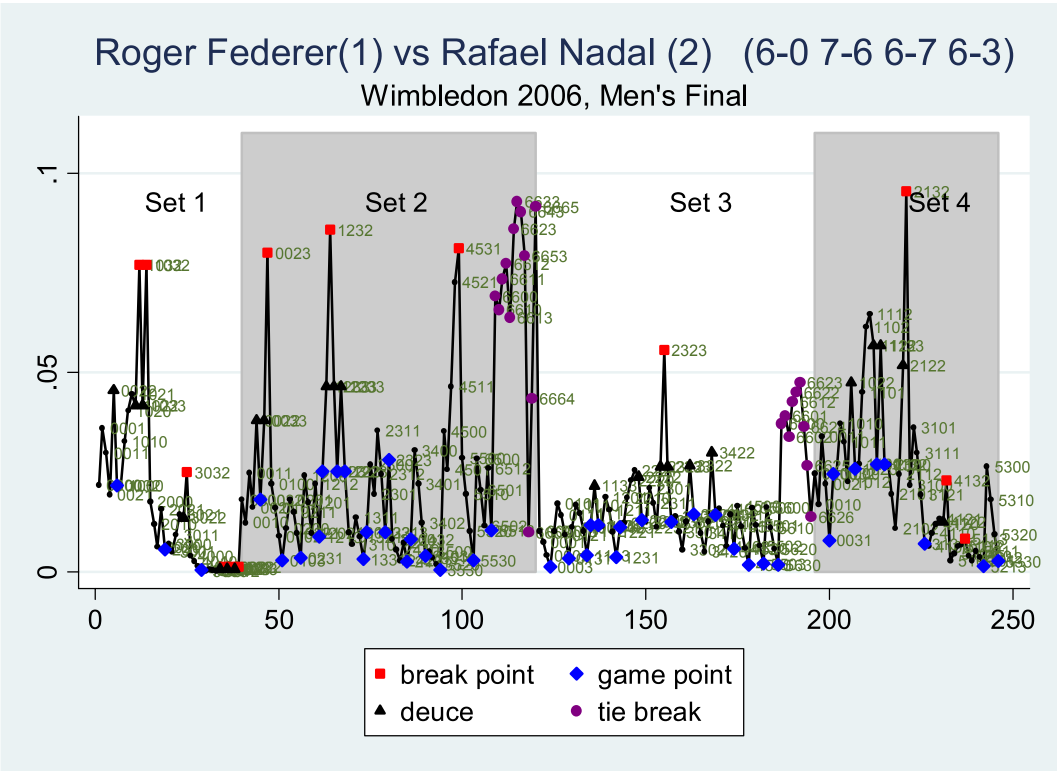
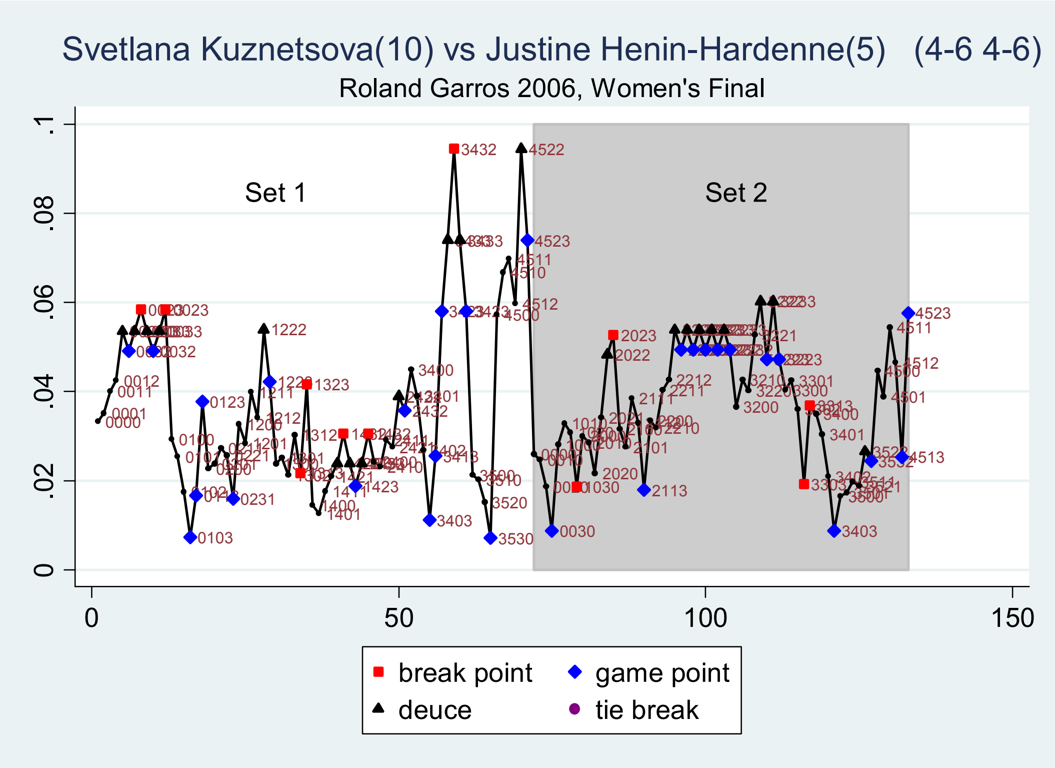


Figure 1: The evolution of importance over the course of selected matches – men

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.



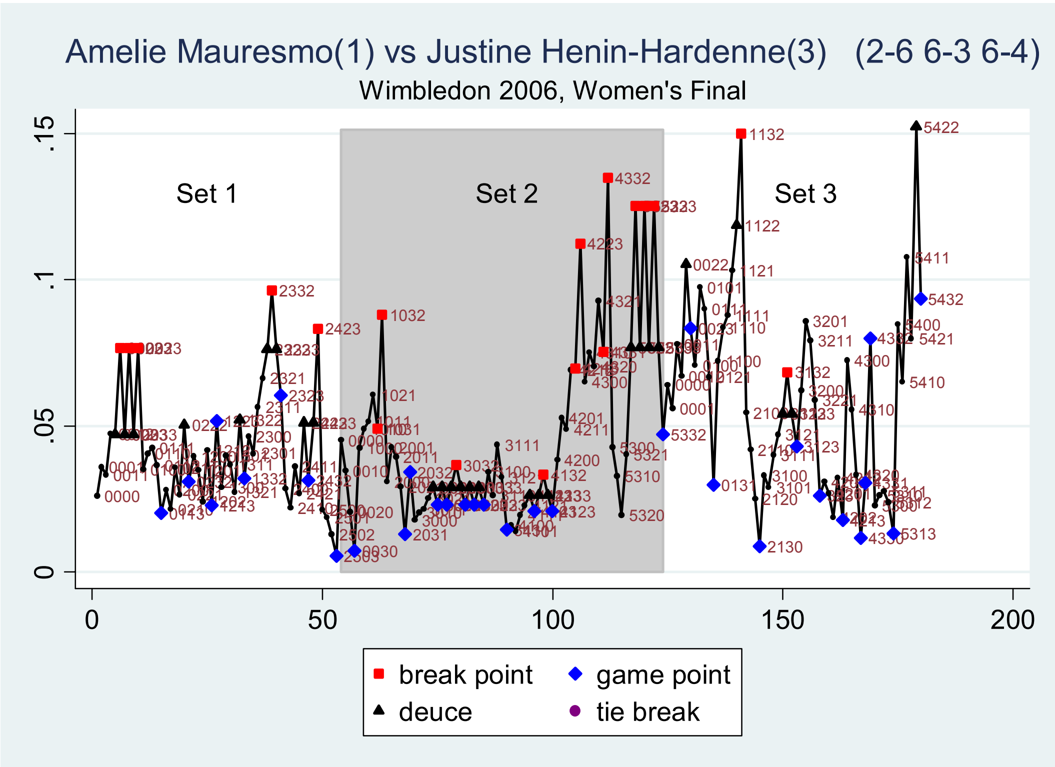


Figure 2: The evolution of importance over the course of selected matches – women

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.

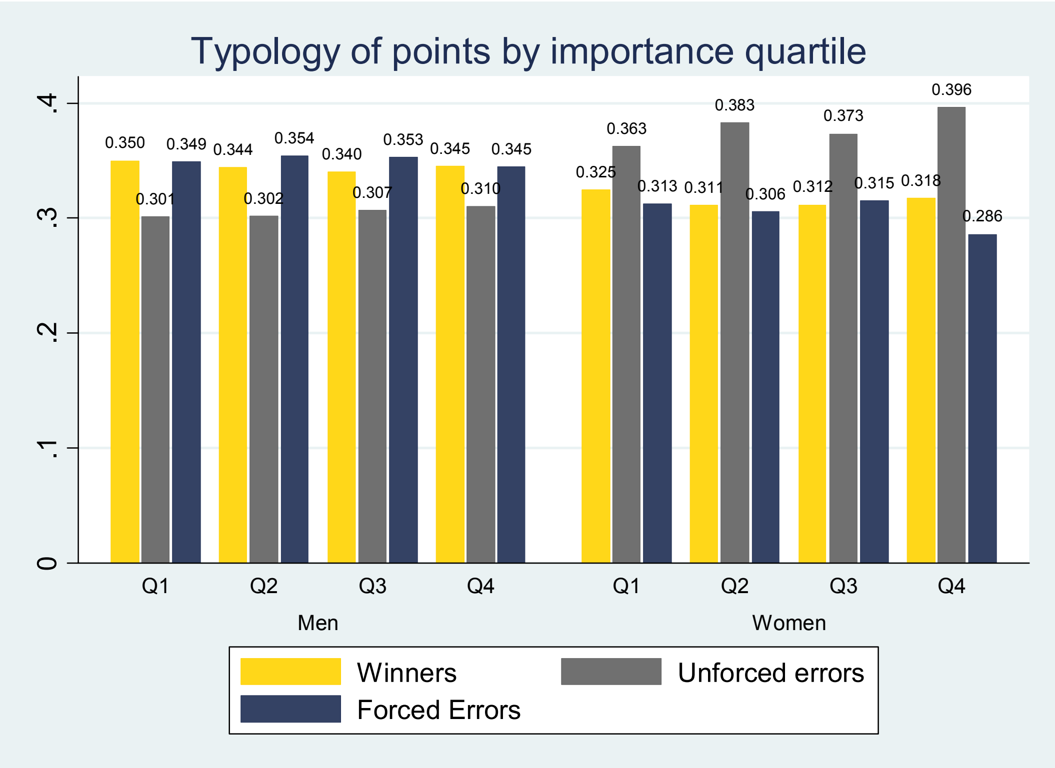


Figure 3: Typology of Points by Importance Quartile

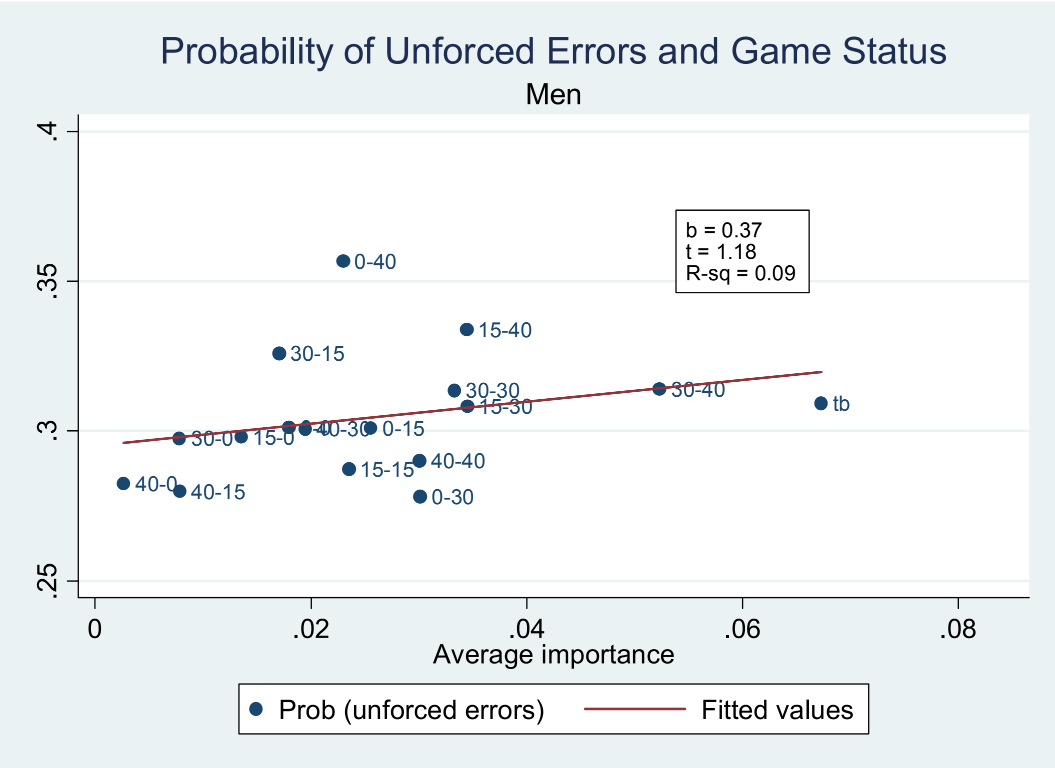


Figure 4a: Probability of Unforced Errors and Game Status – Men

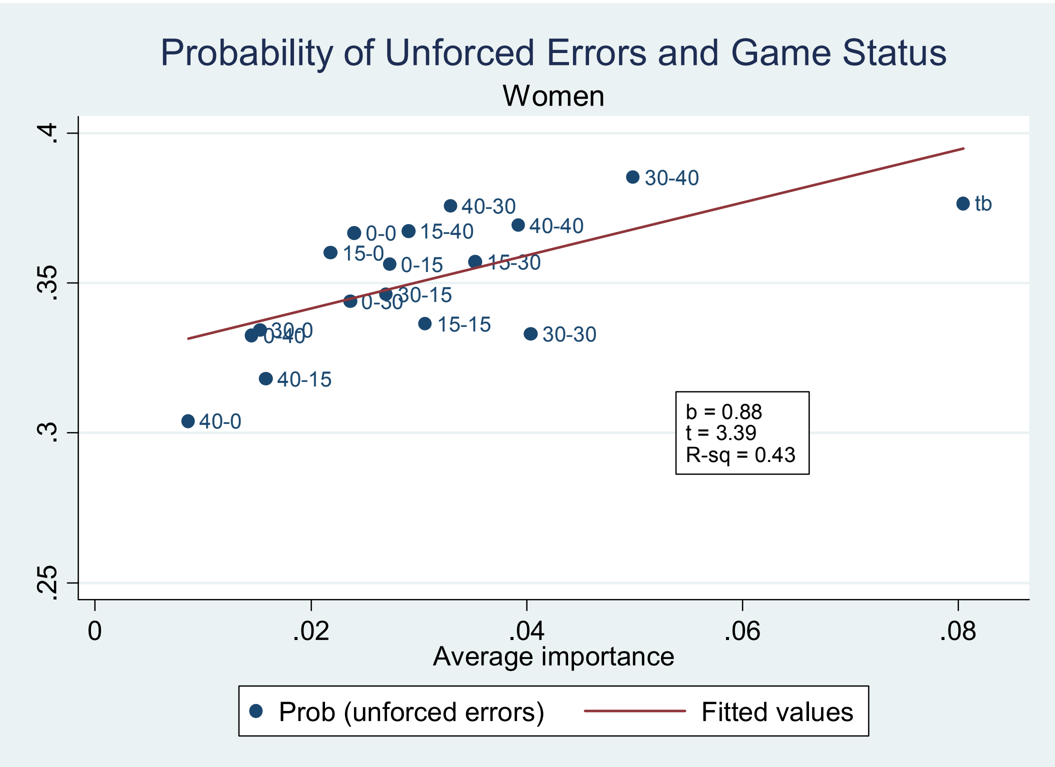


Figure 4b: Probability of Unforced Errors and Game Status – Women

Note: On the vertical axis is the predicted probability of unforced errors estimated from a multinomial logit model for a representative point; on the horizontal axis is the average of the importance value for the particular combination of points won by the server and by the receiver. The fitted line is obtained from the simple regression of the predicted probability on average importance.