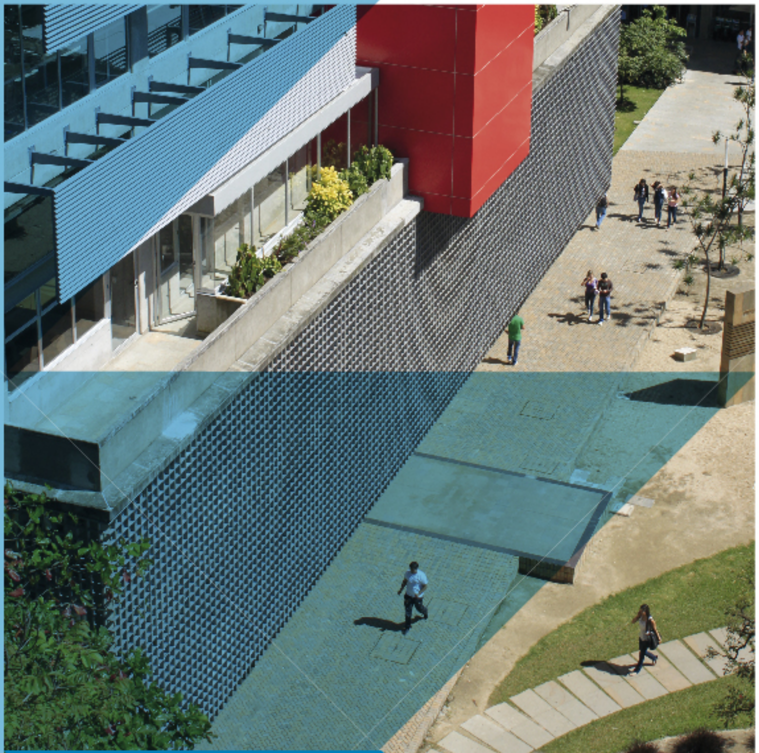
1-2015

Good Girl, Bad Boy: Corrupt Behavior in Professional Tennis

Michael Jetter

Jay K. Walker
Old Dominion University, j1walker@odu.edu

Original Publication Citation
Jetter, Michael and Walker, Jay K., Good Girl, Bad Boy: Corrupt Behavior in Professional Tennis (January 31, 2015). Center for Research in Economics and Finance (CIEF), Working Papers, No. 15-04. Available at SSRN: https://ssrn.com/abstract=2563742 or http://dx.doi.org/10.2139/ssrn.2563742



Escuela de Economía y Finanzas

Documentos de trabajo

Economía y Finanzas

Centro de Investigación
Económicas y Financieras

No. 15-04
2015

Good girl, bad boy: Corrupt behavior in professional tennis

Jetter, Michael; Walker, Jay K.

Good girl, bad boy: Corrupt behavior in professional tennis*

Michael Jetter† Jay K. Walker‡

January 31, 2015

Abstract

This paper identifies matches on the male and female professional tennis tours in which one player faces a high payoff from being "on the bubble" of direct entry into one of the lucrative Grand Slam tournaments, while their opposition does not. Analyzing over 378,000 matches provides strong evidence for corrupt behavior on the men's tour, as bubble players are substantially more likely to beat better ranked opponents when a win is desperately needed. However, we find no such evidence on the women's tour. These results prevail throughout a series of extensions and robustness checks, highlighting gender differences regarding corrupt and unethical behavior, but also concerning collusion. We especially find evidence for collusion once monetary incentives are further increased. Finally, the market for sports betting does not seem to be aware of this phenomenon, suggesting a market imperfection and further confirming our suspicion of irregular activities in men's tennis.

JEL Classification: D73, J16, L83, Z13

Keywords: *cheating, corruption, gender differences, Oaxaca decomposition, sport, tennis*

*We are grateful to Hadi Asiaiefard for excellent research assistance in computing the data downloads from the ATP and WTA websites. We also thank Felix Feldhaus (Head of Latin America at Tipico Co Ltd.) for providing us with data on betting odds from bet365. All remaining errors are our own.

†Department of Economics, Universidad EAFIT, Carrera 49 7 Sur-50, Avenida Las Vegas, Medellín, Colombia; Tel: +57 (317) 433-7003; Email: mjetter7@gmail.com; Web: http://www.michaeljetter.com

‡Department of Economics and Finance, Niagara University, Lewiston, NY, USA; Tel: +1 (716) 286-8311; Email: jwalker@niagara.edu; Web: http://www.jaykwalker.com

1 Introduction

Unethical behavior is difficult to detect. Although experiments and surveys can help us analyze cheating or corrupt behavior, their results remain subject to criticism.[1] Actual labor market outcomes, on the other hand, reveal one's preferences and actions quite clearly. Unfortunately, outcomes in conventional workplaces usually do not allow for identifying corrupt activities, especially if workers know that they are monitored. Ideally a researcher would like to be able to study corrupt behavior in real life without informing workers of the possibility to identify unethical actions.

The following pages analyze corrupt behavior in professional tennis, a sport where men and women compete for comparable payoffs in separated labor markets. Our empirical strategy exploits the unique structure of professional tennis, creating one particular scenario where match fixing between opponents can be beneficial. Specifically, we take advantage of the fact that entering one of the four lucrative Grand Slam tournaments requires a certain ranking. At the entry deadline only the highest ranked 104 players directly qualify for a spot in the main draw, where a player is guaranteed approximately US$30,000 in prize money. As players ranked around #100 and lower are earning at best around US$300,000 per year (before subtracting costs), participation in the four Grand Slam tournaments alone can guarantee a player about 35 to 50 percent of their annual income.[2] Thus, some encounters in tournaments before Grand Slam deadlines are marked by one player desperately needing a win to collect the necessary ranking points, whereas the opponent merely competes for the prize money offered at the current tournament.

Our study draws from a rich sample of over 378,000 total tennis matches on the men's Association of Tennis Professionals (ATP) and the Women's Tennis Association (WTA) Tours,

[1]For instance, an experimental setup in a laboratory may not necessarily produce representative behavior from individuals. Important studies analyzing corrupt behavior in an experimental setting include Fisman and Miguel (2007) and Barr and Serra (2010), who find that cultural origin can predict corrupt behavior. Mocan (2008) and Niehaus and Sukhtankar (2013) provide interesting analyses using micro-level survey data, although Olken (2009) finds corruption perception is not necessarily representative of reality. For summaries of the literature on corruption determinants, one may consider Treisman (2000) or, more recently, Serra (2006).

[2]Numerous popular press articles have highlighted the strong income divide between the absolute top tennis players and players ranked outside the top 100 (most recently Bednall, 2015). It is important to note that prize income is stated before taxes and expenses for traveling, coaches, stringing rackets, etc. The International Tennis Federation (ITF) estimates the annual costs of a full tennis professional around US$160,000 (see Bednall, 2015). Most players compete in about 20 to 35 events per year. See websites of the respective professional tours.

containing all career matches of players ranked in their official listings as of September 21, 2014. Our analysis reveals male players on the bubble are substantially more likely to beat a higher ranked opponent than in other, regular matches (41 percent versus 34 percent). We do not find this anomaly for the female tour. Further evidence comes from the fact that these results become even more pronounced following the 2012 season, when the guaranteed prize money for participants of Grand Slam tournaments increased markedly.

Our findings contribute to two distinct areas of research within economics. First, sport is not exempt from corrupt and unethical behavior. Following seminal papers by Duggan and Levitt (2002) and Wolfers (2006), who study unethical behavior in sumo wrestling and basketball, this paper is, to our knowledge, the first to systematically analyze professional tennis. Second, we add to the existing literature surrounding gender differences in ethical behavior. Isolating gender differences in a conventional workplace environment becomes difficult, as men and women usually compete for the same jobs and work side-by-side. Furthermore, individual outcomes are difficult to observe in many occupational areas, both for lack of clear performance measures and separating individual contributions from team efforts. However, the sport of tennis produces a unique work environment, where these problems are minimized. Most importantly, tennis is a single sport where individual results are easily observed and skewing a desired result becomes easier than in a team environment. In addition, tennis is one of the few sports where men and women earn comparable amounts of compensation, yet compete in separated events.[3]

Although there exists some evidence that women tend to value ethical behavior more so than men, most of these findings are drawn from surveys (e.g., Grove et al., 2011) or aggregate country level data (Swamy et al., 2001; Dollar et al., 2001).[4] However, it is difficult to interpret the results from these analyses as causal, since surveys only exhibit stated preferences, not revealed preferences evaluating actions. Similarly, studies using aggregated data can be subject to endogeneity, as potentially omitted variables and reverse causality can disguise causal inferences. Our study, on the other hand, analyzes revealed preferences (matches on the professional tours) and profits from a random assignment of bubble matches, as matches are assigned by random

[3]Golf may be the only other exception and recently the economics literature has started to exploit these unique work environments for systematic studies. For instance, Brown et al. (2011) analyzes superstar effects when Tiger Woods participates in a tournament.

[4]Niederle, 2014 provides an up-to-date summary of the existing research on gender differences.

drawing at the beginning of a tournament.

The paper proceeds by providing a brief background of professional tennis, focusing on the above described scenarios of bubble matches. Section 3 introduces our sample data and sketches our empirical strategy. Sections 4 and 5 present our main findings. Section 6 considers alternative explanations and Section 7 analyzes whether betting markets are aware of an anomaly in bubble matches. Finally, we conclude with a short discussion in Section 8.

2 Background

The majority of tournaments on the men's professional tennis tour are organized by the ATP (founded in 1972), whereas the best female professionals compete in events organized by the WTA (founded in 1973). Both tours are characterized by tournaments of differing categories of total prize money and earned ranking points. The most prestigious tournaments are the four Grand Slams: the Australian Open, the French Open, Wimbledon, and the US Open. The overall prize money at these four major tournaments is substantial: in 2013, players were paid a total of A$30,000,000 (Australian Open), €22,042,200 (French Open), £22,560,000 (Wimbledon), and US$34,252,000. (Prize money for men and women is equal at these tournaments.) Although these events are technically organized by the International Tennis Federation (ITF, founded in 1913), all ranking points obtained count for the ATP or the WTA tour.

2.1 Structure of Tournaments, Prize Money, and Rankings

Below the Grand Slams, both respective professional tours organize numerous smaller tournaments. For the males these are, in descending order of importance, the ATP tournament categories 1000 (9 tournaments in 2013), 500 (11 tournaments), and 250 (40 tournaments). Below these main ATP tournaments players earn points and prize money at so-called Challengers and Futures tournaments. The lowest group of tournaments (Futures) is officially organized by the ITF and every novice tennis player will begin here to earn their first ranking points in order to move up the rankings. In terms of financial incentives, these events are usually endowed with a total prize money of US$10,000 or US$15,000. Similarly for the women, WTA tournaments of varying importance are organized throughout the year, as well as the smallest events organized by the ITF.

Since prize money and ranking points are increasing with the importance of the tournament, players are naturally seeking to compete in the highest-order events. The four Grand Slam tournaments clearly mark the highest category, where even first round losers earn a paycheck between US$22,538 (Australian Open) and US$36,635 (Wimbledon) as of 2013.[5] If a player manages to win the first round and loses in the subsequent match, s/he earns between US$37,200 (Australian Open) and US$59,200 (Wimbledon). Simply reaching the main draws of all four Grand Slams in 2013 guaranteed a player about US$120,000. As a comparison, the athlete who finished the 2013 season as the 101st best player, Jesse Huta Galung (Netherlands), earned US$107,888 throughout the year. Thus, pure participation in all four Grand Slam tournaments would account for more than his entire annual income.[6] Similarly, the #101 on the WTA Tour, Nadiia Kichenok (Ukraine) earned US$75,913 in prize money throughout 2013.

As a comparison, losing in the first round of the next highest ranked tournament event – the BNP Paribas Open – produces a paycheck of US$11,000, not even one third of the minimum prize money given out at Wimbledon. (Once again, this prize money applies equally to men and women.) Similarly, even *winning* a Challenger tournament on the ATP tour guarantees a payoff of only about US$5,000 to US$15,000 – not even half the amount earned from losing in the first round at Wimbledon. The distribution of prize money in Grand Slams, compared to other tour events, has frequently been discussed in the popular press (e.g., see Oddo, 2013).

In summary, entering the main draw of a Grand Slam tournament is associated with a sizable payday for male and female tennis professionals – a fact that has recently been pointed out in the popular press (Bialik, 2014). However, only 128 players enter the main draw. More precisely, only the best 104 players enter directly with the entry deadline determined six weeks before the main draw of the respective Grand Slam event. Lower ranked players have the chance to compete in a qualifying tournament for 16 additional spots, whereas the remaining 8 spots are

[5] All conversions to US$ from hereon are made at exchange rates taken in January 2015 to facilitate comparability.

[6] Of course, some players also draw income beyond prize money from endorsements, exhibitions, pure prize money tournaments, or league play (especially in Europe). As this information is difficult to gather on a comparable basis, our analysis ignores any incomes beyond prize money on the professional tour. Specifically for players ranked around #100, we are confident additional incomes are comparably low on average. In this context, various press articles highlight the financial struggles of tennis players ranked outside the very top (Morales, 2013; Beaton, 2014; Bednall, 2015).

given out directly by the tournament organizers as "Wild Cards."[7] However, competing in the qualifying tournament entails winning three consecutive matches against virtually even-ranked players for one of the spots in the main draw. For instance, out of the 16 seeds in the men's qualifying draw of the 2013 Australian Open only six eventually qualified. Thus, even though one may be ranked, say, #105, a spot in the lucrative main draw may prove elusive.

2.2 Bubble Matches

Consequently, matches played at tournaments right before the entry deadline are highly important to some players, namely those who are ranked near #104. At these matches, a player ranked around #104 will not only compete for the respective prize money at the current tournament, but also for the ranking points required for the lucrative entry into the upcoming Grand Slam. However, any opponent ranked, say, #50 will only compete for the much smaller prize money at stake in that given tournament. This is precisely where collusion may prove fruitful: the player needing additional ranking points to make the cutoff will have an incentive to offer an otherwise higher ranked (and presumably better) player some reward for letting him win. Note that as draws at all tournaments are produced randomly, bubble matches arise randomly and it is difficult to argue that one can willingly select into the situation of a bubble match.

To get an idea of rankings and points, Table 1 provides an overview of the rankings from #95 to #114 for both tours at the end of the 2013 season. Notice that differences between one ranking spot could be as little as one point. Regarding the points awarded at different rounds for both tours, Tables A1 and A2 provide a detailed overview. For example, a first round win in a Challenger tournament for the males produces between six and ten points, whereas a win in the final of such a tournament would return between 32 and 50 points.

In the following, we label matches where the lower ranked player (the underdog) is ranked between 94 and 114 (ten spots below and above the cutoff) in the two weeks before entry deadlines for the respective Grand Slam tournaments as "bubble matches."[8] This scenario is akin

[7] Wild Cards are usually given to promising young talents who are not ranked highly enough yet or past stars for whom the same applies.

[8] We extend this timeframe to three weeks before the entry deadline to the Australian Open because of the scarce schedule at this point in the season. Since the Australian Open begin mid-January, the entry deadline falls right at the end of the previous season. Our results are robust to changing the cutoff to, for example, including twelve or fifteen spots below or above the cutoff.

to the situation described by Duggan and Levitt (2002) who exploit the fact that sumo wrestlers forego substantial income from finishing a tournament with a losing record. Consequently, some wrestlers have larger incentives to win on the last day of a tournament than others.

One main advantage of our study comes from the fact that tennis is one of the few sports producing a notable tour for female professionals. Specifically, women are able to earn very similar amounts, most famously at the Grand Slam tournaments where prize money at all four events has been equivalent for men and women since 2007 (e.g., see McGregor, 2014). Consequently, the opportunity for collusion arises on both tours and we are able to assess (*i*) whether generally underdogs are more likely to win in those scenarios and (*ii*) whether there exist gender differences in collusive or unethical behavior.

3 Data and Methodology

Our data include information for all career matches of all male and female tennis players who on September 21, 2014, were listed on the respective world rankings (ATP Tour for the men, WTA Tour for the women). On the ATP Tour, this produces 217,153 career matches for 2,025 professionals. On the WTA Tour, our sample consists of 1,261 players and 161,468 matches. Notice that our sample includes all tournaments offering points for world rankings.[9] The women's sample is smaller because female tennis players are usually younger (average of 22.1 years versus 23.8 for the males) and enjoy shorter careers than males.

3.1 Summary Statistics

Table 2 displays summary statistics for all captured matches on both tours. Our dependent variable (Upset) takes on the value of one if the lower ranked player ends up winning the match. Note that we code an upset as the lower ranked player beating a higher ranked opponent, even though we acknowledge that this is not necessarily always considered an upset. In rare scenarios a player could a priori be considered a favorite even though the official rankings might state

[9] The earliest match in our male sample took place on November 23, 1987, whereas the first match of the female sample dates to January 30, 1989. The last match in both samples comes from the week before our data collection (September 21, 2014). For males, this includes all ATP level tournaments (recently 1000, 500, and 250), Challengers, Futures, and Davis Cup. For females, tournaments are categorized into WTA events, ITF tournaments, and the Fed Cup. All data are derived from the ATP and the WTA websites (http://www.atpworldtour.com/ and http://www.wtatennis.com/).

Table 1: Ranks #95 to #114 at the end of the 2013 season.

Rank	Name	Points
Men's Tour		
95	Bedene, Aljaz	573
96	Young, Donald	569
97	Klahn, Bradley	568
98	Stakhovsky, Sergiy	554
99	Falla, Alejandro	552
100	Harrison, Ryan	549
101	Huta Galung, Jesse	549
102	Sock, Jack	545
103	Soeda, Go	543
104	Kavcic, Blaz	542
105	Llodra, Michael	541
106	Hajek, Jan	540
107	Struff, Jan-Lennard	523
108	Klizan, Martin	518
109	Lorenzi, Paolo	515
110	Goffin, David	510
111	Brown, Dustin	506
112	Haider-Maurer, Andreas	506
113	Donskoy, Evgeny	493
114	Kudla, Denis	491
Women's Tour		
95	Pfizenmaier, Dinah	698
96	Vekic, Donna	696
97	Pereira, Teliana	682
98	Giorgi, Camila	671
99	Razzano, Virginie	666
100	Medina Garrigues, Annabel	665
101	Kichenok, Nadiia	643
102	Arruabarrena, Lara	641
103	Duque-Mariño, Mariana	631
104	Petrova, Nadia	628
105	Dolonc, Vesna	627
106	Minella, Mandy	621
107	Martic, Petra	608
108	Pironkova, Tsvetana	600
109	Putintseva, Yulia	597
110	Lucic-Baroni, Mirjana	595
111	Majeric, Tadeja	594
112	Van Uytvanck, Alison	594
113	Vandeweghe, Coco	591
114	Konta, Johanna	586

otherwise. As a famous example, consider grass court events, most notably Wimbledon, that have been notorious for providing unusual playing conditions ("faster" courts compared to other surfaces). Naturally, different conditions favor different styles of play and especially clay court specialists, most prominently from Spain, have in the past been most vulnerable to losing against lower ranked opponents on fast surfaces.[10] However, for the purpose of a consistent analysis, we consider a player as the favorite when facing lower ranked opponents throughout the paper. In our regression analysis we control for potentially confounding features, such as ranking gaps and surfaces.

Table 2: Summary statistics.

	Men's Tour		Women's Tour	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Upset	0.338	(0.473)	0.359	(0.480)
Bubble match	0.002	(0.048)	0.003	(0.057)
Rank of favorite	358	(274)	280	(223)
Rank of underdog	655	(472)	467	(331)
Never played before	0.702	(0.457)	0.751	(0.433)
Never beaten favorite	0.147	(0.354)	0.129	(0.336)
Same country	0.043	(0.203)	0.105	(0.307)
Grand Slam	0.042	(0.201)	0.066	(0.248)
N	217,153		161,468	

In general, we only observe an upset in 33.8 percent of the matches on the ATP Tour – a value that is relatively comparable to that derived from WTA matches (35.9 percent). The second row displays the frequency of what we label as bubble matches: encounters within two weeks prior to an entry deadline to a Grand Slam tournament where the underdog is ranked between #94 and #114.[11] Note that only 0.2 and 0.3 percent of all matches in our sample fall under

[10]For example, consider an article in the New York Times describing the threat by several Spanish players to boycott Wimbledon in 2000 (NYT, 2000).

[11]For the Australian Open we include tournaments held three weeks before the entry deadline since this marks the end of the annual season on both the ATP and the WTA Tour. As the Australian Open, the first of the Grand Slam tournaments in the calendar year, take place mid-January, the entry deadline (6 weeks before the

this category (492 and 518 matches, respectively). In alternative specifications (available upon request), we also tested for different cutoffs for bubble players, such as extending the respective ranking to #124 or excluding players ranked #94 to #103. The derived results are consistent with our main findings.

Table 2 also displays summary statistics of our control variables, such as rankings, whether the two contestants have faced each other on the professional tour before, whether the underdog has beaten the favorite before, and whether both players come from the same country. Since there are less women ranked officially (1,261 as opposed to 2,025 males) at the time of data collection, the average rankings are consequently better for the females. Finally, 4.2 percent of the men's matches are Grand Slam matches, whereas 6.6 percent of women's matches are contested at one of the four major tennis tournaments.

3.2 The Frequency of Upsets in Bubble versus Non-Bubble Matches

Table 3 provides some basic comparisons in terms of the frequency of upsets. Panel A considers the entire sample and we can already spot an interesting difference between the male and female tour. For the males, the probability of an upset is substantially higher in a bubble match, i.e., when an upset is particularly lucrative for the underdog. The frequency of an upset now rises to almost 42 percent, as opposed to 34 percent in regular matches. This difference is strongly relevant in a statistical sense, as indicated by the results from a t-test assuming that the probability of an upset was equally distributed within bubble matches and regular encounters. Column (2) then displays the same statistics for the WTA Tour and we do not find any meaningful anomaly within bubble matches.

Panel B only considers those tournaments taking place in the two weeks before the respective entry deadline for one of the Grand Slam tournaments (three weeks for the Australian Open), labeled as crucial tournaments. Here again, we observe the same pattern: if an underdog is ranked within the critical range of making the main draw the probability of winning rises by almost eight percentage points on the men's tour. Once again, we do not observe this phenomenon for the women's tour.

main draw starts) is scheduled at the end of November. However, there are very few tournaments scheduled for the end of November as the season comes to a close. All our results are robust to just using a two week window for the Australian Open as well.

Table 3: Comparing bubble matches to regular matches.

Variable	Upset Men's Tour	Upset Women's Tour
Panel A: All Matches		
Mean Regular Match	0.338	0.359
N	(216,661)	(160,950)
Mean Bubble Match	0.417	0.363
N	(492)	(518)
T-Test Regular = Bubble Match (p-value)	0.000***	0.862
Panel B: Only Crucial Tournaments		
Mean Regular Match	0.339	0.357
N	(31,923)	(21,893)
Mean Bubble Match	0.417	0.363
N	(492)	(518)
T-Test Regular = Bubble Match (p-value)	0.000***	0.788
Panel C: All Matches		
Mean Regular Tournament	0.338	0.360
N	(184,738)	(139,057)
Mean Crucial Tournament	0.341	0.358
N	(32,415)	(22,411)
T-Test Regular = Crucial Tournament (p-value)	0.340	0.514

Finally, in Panel C we also compare the general probability of an upset in a crucial tournament relative to regular tournaments. After all, it may be possible that these tournaments are especially prone to underdogs beating favorites, possibly for another exogenous reason, such as the previously mentioned particularity of grass court tournaments. Upsets do not seem to be more likely in those tournaments, neither for the males nor the females. Thus, this preliminary statistical evidence suggests those tournaments placed crucially right before the entry deadline to a Grand Slam event are not per se different from other, regular tour events.

However, these basic statistical comparisons could be misleading if other, confounding factors are influencing the probability of an upset. For example, the actual rankings of both players are important as a ranking difference of only a few spots may make for a more equal match-up as opposed to, say, a difference of several hundred spots. To further investigate our research question, we now turn to the results from our regression analyses. Specifically, we present findings from logit regressions estimating the probability of an upset as a function of several regressors, such as ranking parameters, the personal history between both players, surface fixed effects, and time trends. Finally, we apply additional regression techniques introduced throughout the empirical sections.

4 Main Results

Table 4 presents our basic results for the men's (columns 1 through 4) and the women's tour (5 through 8), displaying marginal effects from logit regressions. For either sample, we start by estimating a univariate logit regression where a dummy variable for a bubble match is used to predict the probability of an upset, i.e., a win by the lower ranked player. Indeed bubble matches are 7.5 percent more likely to see an upset on the men's tour. This result is statistically relevant at the one percent level. However, the regression analyzing the female sample (column 5) produces a coefficient far from conventional levels of statistical significance. In terms of magnitude, the estimate only produces about one twentieth of the coefficient generated by the regression using the male data.

For the men, columns (2) through (4) display coefficients gradually adding control variables. We first control for the ranking of the favorite (linear and squared) and the difference in ranking

Table 4: Predicting the probability of an upset, i.e. the lower ranked player beating the higher ranked player. Columns (1) – (4) consider the men's tour, whereas columns (5) – (8) evaluate the women's tour. Displaying marginal effects.

	Men's Tour				Women's Tour			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bubble match	0.075***	0.052***	0.053***	0.048**	0.004	0.004	0.004	-0.003
	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)	(0.020)	(0.020)	(0.020)
Ranking of favorite		0.000***	0.000***	0.000***		0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
(Rank of favorite)^2		-0.000***	-0.000***	-0.000***		-0.000***	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Ranking gap		-0.000***	-0.000***	-0.000***		-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Never played before			-0.041***	-0.044***			-0.048***	-0.048***
			(0.003)	(0.003)			(0.004)	(0.004)
Never beaten favorite			-0.072***	-0.073***			-0.089***	-0.089***
			(0.004)	(0.004)			(0.005)	(0.005)
Same country			-0.013***	-0.011**			-0.016***	-0.016***
			(0.005)	(0.005)			(0.004)	(0.004)
Grand Slam				-0.046***				-0.035***
				(0.005)				(0.005)
Surface FE, time trend (linear and squared)				yes				yes
N	217,153	217,153	217,153	217,091	161,468	161,468	161,468	161,431
Log lik.	-138,936.62	-133,911.92	-133,702.50	-133,567.81	-105,437.76	-102,561.29	-102,369.51	-102,320.13

Notes: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

between both players. Note that the coefficient on bubble matches decreases by almost one third, yet retains its statistical power. Further, column (3) adds categorical variables for whether (*i*) both players have faced each other before, (*ii*) the underdog has beaten the favorite before, and (*iii*) both players compete for the same country. Finally, column (4) controls for a dummy indicating Grand Slam tournaments, surface fixed effects (carpet, clay, and grass with hard courts forming the reference category), and time trends (linear and squared). In this most complete estimation, bubble matches are 4.8 percent more likely to end with an upset compared to any other matches on the men's tour.

For the women, on the other hand, we find no evidence of bubble matches being any more likely to result in an upset than regular matches. Adding the same set of control variables, columns (6) through (8) continue to produce a relationship that is not statistically different from zero. Thus, there seem to be systematic differences in the result of bubble matches between the male and the female tennis tour.

Finally, we want to briefly discuss the implications of the associated control variables. In fact, all remaining regressors return almost identical conclusions for the male and the female sample. For instance, upsets are less likely in Grand Slam tournaments – a result that seems fairly intuitive. As the winner in Grand Slam matches is determined by winning three sets (best-of-five), in regular tournaments one only needs to win two sets (best-of-three). Thus, favorites are more likely to emerge victoriously when given more time to perform, whereas winning "only" two sets is easier for an underdog, as opposed to having to win three sets. Further, if the underdog has never beaten their opponent before (at least on the professional tour) an upset becomes substantially less likely. In the following analysis, we only display the relevant coefficients to facilitate readability, yet all results of the remaining correlates confirm the results from Table 4 (available upon request). With these results in hand, we now move to extensions, alternative specifications, and robustness checks.

5 Extensions and Robustness Checks

So far, we have provided evidence of an upset in a bubble match being systematically more likely than in a regular match on the men's tour, whereas this is not the case on the women's tour. The interpretation of this finding may lead one to believe male players could collude in a bubble match situation or, in other words, the favorite may let the underdog win if stakes are substantially increased for his opponent. Note the importance of a bubble match only applies to *one* player – the underdog. For the favorite, the official payoff of winning the match is purely the respective prize money at the current tournament – a reward that is substantially lower than the ticket into the main draw of the upcoming Grand Slam (the looming reward for the underdog). This section now considers several extensions and alternative specifications of this hypothesis.

5.1 The Likelihood of a Tiebreak

Another, more nuanced way to test whether the outcomes in bubble matches are in any way different from regular matches is to analyze specific scores. In particular, non-Grand Slam matches are played in a best-of-three sets format, where a player wins a set when reaching six games with a difference of two games (for example, 6-4). However, if a set reaches a score of 6-6 a so-called tiebreak is played, where the first to seven points wins (with a difference of two points). This particular form of competition is known to be a very close encounter where every point counts. Now, if we were to assume there exists some sort of arrangement between both players before a match, then we would expect tiebreaks, the closest form of deciding the winner of a set, to be less prevalent than in other, regular matches.

Table 5 displays the results from logit regressions estimating the likelihood of at least one tiebreak occurring in a match. Again, columns (1) through (4) consider the male sample and (5) through (8) analyze the female sample. The first two columns consider all non-Grand Slam matches of our respective samples, including the dummy variable for bubble matches, a binary indicator for upsets, and an interaction term between the two.[12] The intuition here is to specifically evaluate the upsets in bubble matches where match fixing may have potentially occurred. Indeed, we find tiebreaks to be rare (12.8 percent less likely) in those bubble matches where the underdog prevails. This result virtually remains unchanged when including all control variables presented in Table 4. Note that we exclude about 7.7 percent of our sample matches

[12]As Grand Slam matches are competed in a best-of-five set format, we exclude them from this particular analysis.

that have been terminated either by injury, withdrawal, or default. As for the women's tour (columns 5 and 6), we find no evidence of tiebreaks being less likely in those bubble matches that eventually result in an upset.

As a further robustness check, we then only analyze the upsets for either sample. Here again, upsets in bubble matches on the men's tour are less likely to witness a tiebreak – a result that is statistically relevant on the one percent level. In terms of magnitude, upsets in bubble matches are almost 10 percent less likely to exhibit at least one tiebreak. As before, we find no such evidence on the WTA Tour. This strengthens the argument of arrangements taking place in a non-trivial amount of bubble matches.

5.2 Are Bubble Matches Different in Other Aspects?

The fact that upsets are more likely in bubble matches does not necessarily mean collusion between both players occurs. For instance, bubble matches could systematically differ from regular matches in other characteristics – an explanation that is difficult to filter out in generic logit regressions. Specifically, the relationship between any other control variable and the probability of an upset is by assumption identical in bubble matches and in regular matches in our main estimation framework.

However, it is possible to technically distinguish between two sorts of observations within a sample by using the Oaxaca-Blinder decomposition (see Oaxaca, 1973, and Blinder, 1973; also see Grove et al., 2011, for a recent application). Although traditionally used to trace out differences across gender or race in wage estimations, the technique can be applied to any (larger) sample where we can exogenously distinguish between two different types of observations. Given the focus of our paper, it comes naturally to distinguish between regular and bubble matches.

As a result, the Oaxaca-Blinder decomposition produces two main coefficients explaining the difference between the probability of an upset in bubble versus regular matches: one that displays how much is explained by endowments (other characteristics, such as rankings) and another that explains how much is explained by coefficients (i.e., purely the fact of one match being a bubble match).[13]

[13]In Stata, we apply the Oaxaca command, as introduced by Jann (2008). Using a more refined econometric framework for estimating binary outcome variables, provided by Sinning et al. (2008), produces very similar results (about 75 percent of the difference in upset probability being explained by coefficients, rather than characteristics).

Table 5: Results from Logit regressions predicting the probability of at least one tie-break in the match.

	Men's Tour				Women's Tour			
	All Non-Grand Slam Matches		All Upsets in Non-Grand Slam Matches		All Non-Grand Slam Matches		All Upsets in Non-Grand Slam Matches	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bubble match	0.066***	0.025	-0.069*	-0.096***	0.002	-0.013	0.042	0.022
	(0.025)	(0.025)	(0.035)	(0.035)	(0.021)	(0.021)	(0.027)	(0.027)
Upset	0.070***	0.060***			0.039***	0.035***		
	(0.002)	(0.002)			(0.002)	(0.002)		
Bubble match × Upset	-0.128***	-0.118***			0.036	0.038		
	(0.041)	(0.041)			(0.032)	(0.032)		
Control variables[a]		yes		yes		yes		yes
N	200,416	200,356	66,947	66,934	146,957	146,920	52,498	52,492

Notes: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
[a]Includes ranking of the favorite (linear and squared), ranking gap, never played before, never beaten favorite, same country, surface FE, time trend (linear and squared).

Table 6 displays the results subsequently incorporating our control variables. Overall, the difference between the probability of an upset in a regular match versus a bubble match comes out to be about 7.9 percentage points (41.7 minus 33.8 percent). As seen previously, that difference is statistically powerful, indicating that we observe a much higher likelihood of an upset in matches where the underdog is on the bubble. Note that coefficients are able to explain between 68 and 75 percent of that difference, depending on the set of control variables used (0.054 or 0.059 of 0.079). Thus, bubble matches are more likely to end in an upset purely because of their status as bubble matches, but not because of other surrounding characteristics, as endowments are not statistically relevant.

5.3 Relating Bubble Matches to Prize Money

Until now, we have analyzed a general difference between bubble matches and regular tour matches. However, the prize money structure of Grand Slam tournaments changed substantially in recent years. Although reaching the main draw of a Grand Slam has long been a milestone for obtaining substantial cash rewards, there have been sizable recent increases in monetary payments (for instance, consider Oddo, 2013, who highlights these changes are unique to Grand Slam tournaments). In fact, the Players Associations (ATP and WTA) agreed on not only a major raise in prize money at Grand Slam tournaments, but also on a more even distribution towards those players who lose early. Specifically, consider a comparison of the prize money awarded to a first-round loser in each of the four Grand Slam events in the past five years, displayed in Figure 1.

Although the minimal prize money from reaching the main draw has been high previously relative to other tournaments, there have been major changes in prize money after 2012 for at least three of the four major events. The Australian Open, Wimbledon, and the US Open each increased the sum awarded to a first-round loser by 33, 62, and 39 percent, respectively. If prize money at the elusive Grand Slam tournaments was indeed a motivation to collude with an opponent, we would expect to see more upsets in bubble matches after the prize money increased significantly.

As a reference point and baseline in all decompositions, we select regular matches, as these constitute the vast majority of our samples.

Table 6: Results from Oaxaca-Blinder decompositions for the men's tour, predicting the probability of an upset.

	(1)	(2)	(3)	(4)
Regular Match	0.338***	0.338***	0.338***	0.338***
	(0.001)	(0.001)	(0.001)	(0.001)
Bubble Match	0.417***	0.417***	0.417***	0.417***
	(0.022)	(0.022)	(0.022)	(0.022)
Difference	-0.079***	-0.079***	-0.079***	-0.079***
	(0.022)	(0.022)	(0.022)	(0.023)
Endowments	3.093	3.280	3.476	3.122
	(4.947)	(4.987)	(5.013)	(5.062)
Coefficients	-0.059***	-0.059***	-0.054**	-0.054**
	(0.022)	(0.022)	(0.022)	(0.022)
Interaction	-3.113	-3.299	-3.500	-3.146
	(4.947)	(4.987)	(5.013)	(5.062)
Control Variables				
Ranking[a]	yes	yes	yes	yes
Head-2-Head[b]		yes	yes	yes
Grand Slam, Surface FE			yes	yes
Time trend (linear & squared)				yes
N	217,153	217,153	217,091	217,091

Notes: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
[a]Includes rank of favorite, (rank of favorite)^2, and ranking gap.
[b]Includes never played before, never beaten before, and same country.

Figure 1: Prize Money for 1st Round Losers in Grand Slam Tournaments (local currencies).

Table 7 provides basic summary statistics when splitting our samples into before and after the entry deadline for the 2013 Australian Open, when the increase in prize money occurred. Interestingly, upsets in bubble matches were more likely on the men's tour before (about 40 percent versus 34.3 percent), but recently this trend has become even more pronounced. In fact, after the increase in prize money at Grand Slam tournaments nearly 50 percent of all bubble matches end in an upset. Thus, since the 2013 Australian Open entry deadline we see upsets in about 32 percent of all matches, but in nearly every other bubble match.

Table 7: Probability of upsets before and after the entry deadline for the 2013 Australian Open.

Variable	Upset Men's Tour	Upset Women's Tour
Panel A: All Matches Before AO 2013 Entry Deadline		
Mean Regular Match	0.343	0.360
N	(167,068)	(128,823)
Mean Bubble Match	0.398	0.360
N	(379)	(425)
T-Test Regular = Bubble Match (p-value)	0.024**	0.988
Panel B: All Matches Since AO 2013 Entry Deadline		
Mean Regular Match	0.320	0.355
N	(49,593)	(32,127)
Mean Bubble Match	0.478	0.376
N	(113)	(93)
T-Test Regular = Bubble Match (p-value)	0.000***	0.666

We now test these statistical irregularities in a logit regression framework, controlling for potentially confounding factors. The results displayed in Table 8 further confirm our suspicion: upsets in bubble matches on the men's tour are especially prevalent after prize money increases at the beginning of the 2013 season. In fact, before that change the respective coefficient is not significant on conventional levels, although still somewhat sizable at about 0.03. Remember that the respective coefficient from the baseline regression in Table 4 reaches a magnitude of 0.048. Since the 2013 Australian Open, however, that coefficient increases markedly to 0.1, meaning a bubble match is ten percent more likely to produce an upset than any other match. Once again, we observe none of these dynamics for the female data in columns (3) and (4). These results lend further support to the idea that male tennis players tend to collude and fix matches, as the incidence of bubble matches ending in an upset increases when the stakes are higher.

Table 8: Predicting the probability of an upset on the men's tour, i.e., the lower ranked player beating the higher ranked player. Displaying marginal effects.

	Men's Tour		Women's Tour	
	(1) Since AO 2013	(2) Before AO 2013	(3) Since AO 2013	(4) Before AO 2013
Bubble Match	0.101^{**}	0.030	0.032	0.020
	(0.040)	(0.023)	(0.048)	(0.023)
Control variables^{a}	yes	yes	yes	yes
N	49,668	167,423	32,220	129,211
Log lik.	-29,789.56	-103,791.35	-20,868.98	-84,010.516

Notes: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
[a]Includes rank of favorite, (rank of favorite)^2, ranking gap, never played before never beaten before, same country, Grand Slam, and surface FE.

6 Alternative Explanations

We now consider two natural explanations for why an underdog may win an ATP match carrying such far-reaching consequences for him. First, the match could hold lower marginal benefit for the favorite and he may choose to not compete fully, potentially saving his energy. Second, it may be possible that the underdog places more effort into this specific match. For instance, he may choose to prepare more extensively for this particular tournament knowing of its importance or compete harder. The following discussion addresses both explanations.

6.1 Are the Losers Under-performing?

First, we form a subsample identifying those players who at some point in their career lost to a player on the bubble. These players are the potential recipients of match fixing, presumably letting the lower ranked opponent win for a reward. We then select all career matches where those players have been the favorite. Specifically, we wish to check whether these players are just particularly prone to upsets in comparable matches. For instance, it could be that those players only compete heavily if a big reward is on the line, such as in Grand Slam events. In smaller tournaments, on the other hand, they may choose to put in a lower level of effort. If this were the case, our prior findings would be spurious.

Table 9 displays results from logit regression estimating the probability of an upset, where we once again include a dummy variable for bubble matches. We find little evidence for this subsample of players to be generally more prone to losing to underdogs, as an upset remains substantially more likely in a bubble match. This result is robust to the inclusion of all control variables, such as ranking characteristics, head-to-head history, surface fixed effects, and time trends. Column (6) then also includes Grand Slam matches and the coefficient of interest remains virtually unchanged. As before, Grand Slam matches are less likely to witness an upset, consistent with our intuition. These results show it is the particular characteristic of a bubble match for the lower ranked opponent driving the upset and not an intrinsic feature of the favorite who could be prone to losing to underdogs from time to time. This further lends evidence to the argument of some favorites letting their opponent win in what constitutes an important match for that opponent.

6.2 Are the Winners Over-performing?

Second, it may be possible that the player on the bubble puts in an extra effort. This explanation comes as fairly intuitive, as higher payoffs have been shown to provoke better effort (Lazear, 2000). Thus, we now derive a subsample that (*i*) includes only those players who at some point won a match against a higher ranked opponent in what constitutes a bubble match for them; and (*ii*) incorporates high-stakes matches (i.e., where similar amounts of prize money are at stake) where the respective player has been an underdog. In particular, we include Grand Slam matches in which every match won in the main draw increases the previous prize money by

Table 9: Considering the performance of ATP players who at some point in their career lost to a lower ranked opponent in what constitutes a bubble match for their opponent. This subsample excludes Grand Slams with the exception of column (6) that incorporates all their matches. Displaying marginal effects.

	(1)	(2)	(3)	(4)	(5)	(6)
Bubble Match	0.171***	0.129***	0.131***	0.130***	0.130***	0.132***
	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Ranking[a]	yes	yes	yes	yes	yes	yes
Head-2-Head[b]			yes	yes	yes	yes
Surface FE				yes	yes	yes
Time trend (linear and squared)					yes	yes
Grand Slam						-0.056***
						(0.008)
N	38,856	38,856	38,856	38,850	38,850	43,184
Log lik.	-24,375.59	-23,711.33	-23,668.67	-23,664.28	-23,663.00	-26,206.47

Notes: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
[a]Includes rank of favorite, (rank of favorite)^2, ranking gap, and same country.
[b]Includes never played before, never beaten before, and same country.

about 60 to 80 percent.[14] Thus, we could reasonably assume that if a player puts in an extra amount of effort in a bubble match he would try to reach that level of effort in a Grand Slam match as well, given that the monetary reward from both matches is comparably high. Keep in mind that winning a round in a Grand Slam tournament (or similarly reaching the main draw of a Grand Slam tournament) *guarantees* one the prize money of the next round, but also provides the opportunity to win another match and further increase prize money.

Table 10 displays the results from logit regressions, evaluating whether those players who beat higher ranked opponents when on the bubble are generally more prone to upsetting favorites. Columns (1) through (5) consider bubble matches and all Grand Slam matches, where payoffs are comparable (or higher) than in bubble matches. The importance of bubble matches reveals these players are not usually prone to produce upsets – this is only the case in bubble matches. Thus, an increased effort level by the underdogs in bubble matches is unlikely to solely explain our findings, since we do not observe such patterns in matches of similar importance. Further, column (6) evaluates all career matches of one-time winners of bubble matches (not only high-stakes encounters) and the coefficient of interest remains virtually unchanged.

7 Does the Market Know?

After evaluating several alternative explanations, we now move to a market that is closely associated with professional sports: the betting market. In the context of our narrative, a natural question to emerge is whether the market is aware of bubble matches being more likely to produce upsets. To test this, we access data for 15,292 sample matches offered by bet365, one of the world's leading online sports betting companies. To evaluate whether bubble matches are in any way standing out in these odds, we regress the odds offered for an upset on a dummy variable for bubble matches and then add further controls. The results are displayed in Table 11.

In a univariate framework, bubble matches do indeed receive a higher likelihood of an upset, indicating a general awareness of our discovered phenomenon. However, once rankings and ranking gaps are added in column (2), the statistical role of bubble matches disappears completely.

[14] For instance, in the 2013 Australian Open, the prize money for a second round loser was A$45,500 – almost A$20,000 more than for a first round loser.

Table 10: Considering the performance of ATP players who at some point in their career beat a better ranked opponent in a bubble match. This subsample includes bubble matches and Grand Slam matches. Column (6) considers all tour matches by those players. Displaying marginal effects..

	(1)	(2)	(3)	(4)	(5)	(6)
Bubble Match	0.248***	0.176***	0.180***	0.169***	0.168***	0.170***
	(0.023)	(0.024)	(0.024)	(0.025)	(0.025)	(0.025)
Ranking[a]		yes	yes	yes	yes	yes
Head-2-Head[b]			yes	yes	yes	yes
Surface FE				yes	yes	yes
Time trend (linear and squared)					yes	yes
Grand Slam						-0.078***
						(0.011)
N	2,760	2,760	2,760	2,758	2,758	23,402
Log lik.	-1703.79	-1599.17	-1593.15	-1591.68	-1588.84	-15342.82

Notes: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
[a]Includes rank of favorite, (rank of favorite)^2, and ranking gap.
[b]Includes never played before, never beaten before, and same country.

Table 11: OLS regressions estimating the probability given to an upset, as evaluated by the betting markets (Bet365).

	(1)	(2)	(3)	(4)	(5)	(6) Since AO 2013
Bubble match	0.048***	-0.008	-0.003	-0.008	-0.009	0.007
	(0.012)	(0.010)	(0.010)	(0.010)	(0.010)	(0.020)
Ranking[a]		yes	yes	yes	yes	yes
Head-2-Head[b]			yes	yes	yes	yes
Surface FE, Grand Slam				yes	yes	yes
Time trend (linear and squared)					yes	yes
N	15,292	15,292	15,292	15,292	15,292	2,867
R^2	0.001	0.260	0.297	0.307	0.309	0.372

Notes: Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[a]Includes rank of favorite, (rank of favorite)2, ranking gap, and same country.
[b]Includes never played before, never beaten before, and same country.

Columns (3) to (5) then add the remaining independent variables that may affect the odds given to an upset (head-to-head characteristics, surface fixed effects, a dummy for Grand Slams, and time trends), but the effect from bubble matches does not differ from zero in a statistical sense. Thus, betting markets do not seem to anticipate the higher likelihood of upsets in bubble matches.

8 Conclusion

This paper analyzes professional tennis matches producing a unique payoff scenario where one player can obtain a much larger monetary benefit from winning a match than their opponent. Since only 104 players enter directly into each of the four lucrative Grand Slam tournaments, being ranked #104 or better right before the entry deadline becomes a crucial goal for players, on the men's as well as on the women's tour (ATP and WTA). Thus, in matches preceding the entry deadline the payoff for some players becomes large, exceeding way beyond the contested prize money at regular tournaments. Our analysis then focuses on those matches where a player "on the bubble" is facing a better ranked opponent right before the entry deadline closes.

We find upsets (beating a higher ranked opponent) to be substantially more likely in bubble matches on the men's tour, but *not* on the women's tour. This finding remains consistent in all estimated specifications, suggesting that upsets in bubble matches on the men's tour are about five percent more likely than in an ordinary match. These findings remain robust to controlling for several potentially confounding factors, such as the characteristics associated with rankings, the personal head-to-head history, the tournament category, surface fixed effects, and time trends. In addition, the likelihood of a tiebreak (the closest form of determining the winner of a set) occurring in a match is substantially lower in those matches where the player on the bubble needs the upset. Further, the effect becomes more notable after prize money for participating in Grand Slam events increased substantially with the beginning of the 2013 season. Since then, underdogs have won almost 50 percent of bubble matches, as opposed to a general likelihood of 32 percent in non-bubble matches. This further confirms our suspicion of corrupt behavior taking place on the men's professional tennis tour.

Finally, we find betting agencies are systematically unaware of this phenomenon, as an upset in a bubble match is not specifically predicted by the market. This only further strengthens the idea of a systematic irregularity in male tennis and provides evidence for match fixing. Overall, our results provide evidence for systematic corrupt behavior on the men's tennis tour, but not the women's. Beyond the obvious implications for the sport of tennis, this further contributes to the large literature on gender differences in corrupt behavior, collusion, and cheating.

References

Barr, A. and Serra, D. (2010). Corruption and culture: An experimental analysis. *Journal of Public Economics*, 94(11):862–869.

Beaton, A. (2014). Top 32 highest-paid players in several sports. *Wall Street Journal*, August 26, 2014.

Bednall, J. (2015). Gap between the haves and have nots in tennis continues to widen. *News.com.au*, January 29, 2015.

Bialik, C. (2014). Tennis has an income inequality problem. *FiveThirtyEight*, December 30, 2014.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4):436–455.

Brown, J. et al. (2011). Quitters never win: the (adverse) incentive effects of competing with superstars. *Journal of Political Economy*, 119(5):982–1013.

Dollar, D., Fisman, R., and Gatti, R. (2001). Are women really the "fairer" sex? Corruption and women in government. *Journal of Economic Behavior & Organization*, 46(4):423–429.

Duggan, M. and Levitt, S. D. (2002). Winning isn't everything: Corruption in sumo wrestling. *American Economic Review*, 92(5):1594–1605.

Fisman, R. and Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy*, 115(6):1020–1048.

Grove, W. A., Hussey, A., and Jetter, M. (2011). The gender pay gap beyond human capital: Heterogeneity in noncognitive skills and in labor market tastes. *Journal of Human Resources*, 46(4):827–874.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4):453–479.

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

McGregor, J. (2014). How women's tennis fought for equal pay. *Washington Post*, September 5, 2014.

Mocan, N. (2008). What determines corruption? International evidence from microdata. *Economic Inquiry*, 46(4):493–510.

Morales, M. (2013). How the 92nd-ranked tennis player in the world earns a comfortable living. *Forbes*, August 26, 2013.

Niederle, M. (2014). Gender. *NBER Working Paper 20788.*

Niehaus, P. and Sukhtankar, S. (2013). Corruption dynamics: The golden goose effect. *American Economic Journal: Economic Policy*, 5(4):230–269.

NYT (2000). Plus: Tennis – Wimbledon; Spain's top players threaten to boycott. *New York Times*, June 20, 2000.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709.

Oddo, C. (2013). Prize money rises for slams; not so much elsewhere. *USA Today*, August 28, 2013.

Olken, B. A. (2009). Corruption perceptions vs. corruption reality. *Journal of Public Economics*, 93(7):950–964.

Serra, D. (2006). Empirical determinants of corruption: A sensitivity analysis. *Public Choice*, 126(1-2):225–256.

Sinning, M., Hahn, M., and Bauer, T. K. (2008). The Blinder-Oaxaca decomposition for non-linear regression models. *Stata Journal*, 8(4):480–492.

Swamy, A., Knack, S., Lee, Y., and Azfar, O. (2001). Gender and corruption. *Journal of Development Economics*, 64(1):25.

Treisman, D. (2000). The causes of corruption: A cross-national study. *Journal of Public Economics*, 76(3):399–457.

Wolfers, J. (2006). Point shaving: Corruption in NCAA basketball. *American Economic Review*, 96(2):279–283.

Appendix

Table A1: Distribution of ranking points on the male Tour (ATP), as of 2014. Source: ATP Points Breakdown.

Round[a]	W	F	SF	QF	R16	R32	R64	R128	Q
Grand Slams	2000	1200	720	360	180	90	45	10	25
Barclays ATP World Tour Finals	*1500								
ATP World Tour Masters 1000	1000	600	360	180	90	45	10(25)	(10)	(1)25
ATP 500	500	300	180	90	45	(20)			(2)20
ATP 250	250	150	90	45	20	(5)			(3)12
Challenger 125,000 +H[b]	125	75	45	25	10				5
Challenger 125,000	110	65	40	20	9				5
Challenger 100,000	100	60	35	18	8				5
Challenger 75,000	90	55	33	17	8				5
Challenger 50,000	80	48	29	15	7				3
Challenger 40,000 +H[b]	80	48	29	15	6				3
Futures 15,000 +H[b]	35	20	10	4	1				
Futures 15,000	27	15	8	3	1				
Futures 10,000	18	10	6	2	1				

Notes: *Barclays ATP World Tour finals 1500 for undefeated Champion (200 for each round robin match win, plus 400 for a semi-final win, plus 500 for the final win).
(1) 12 points only if the main draw is larger than 56.
(2) 10 points only if the main draw is larger than 32.
(3) 5 points only if the main draw is larger than 32.
[a]W = Winner; F = Final; SF = Semifinal; QF = Quarterfinal; Q = Qualifier.
[b]H indicates that Hospitality is provided.

Table A2: Distribution of ranking points on the female Tour (WTA), as of 2014. Source: 2014 WTA Ranking System.

Round[a]	W	F	SF	QF	R16	R32	R64	R128	QLFR	Q3	Q2	Q1
Grand Slams	2000	1300	780	430	240	130	70	10	40	30	20	2
BNP Paribas WTA Championship	1500*	1050*	690*	(+70 per Round Robin Match; 160 per Round Robin Win)								
Tournament of Champions Sofia	375*	255*	180*	(+25 per Round Robin Match; 35 per Round Robin Win)								
Premier Mandatory: Beijing, Indian Wells, Madrid, Miami												
96 Singles (48Q)	1000	650	390	215	120	65	35	10	30		20	2
64/60 Singles (32Q)	1000	650	390	215	120	65	10		30		20	2
Premier 5 (5): Cincinnati, Doha, Montreal, Rome, Wuhan												
56 Singles (64 Q)	900	585	350	190	105	60	1		30	22		
56 Singles (48/32 Q)	900	585	350	190	105	60	1		30	20		
Premier (12):												
56 Singles	470	305	185	100	55	30	1		25		13	1
32 Singles	470	305	185	100	55	1			25	18	13	1
International Events (32):												
32 Singles (32Q)	280	180	110	60	30	1			18	14	10	1
32 Singles (16Q)	280	180	110	60	30	1			18		12	1
WTA 125K Series	160	95	57	29	15	1			6		4	1
ITF Circuit Events:												
ITF$100,000+H (32/16)	150	90	55	28	14/1	1			6	4	1	
ITF $100,000 (32/16)	140	85	50	25	13/1	1			6	4	1	
ITF $75,000+H (32/16)	130	80	48	24	12/1	1			5	3	1	
ITF $75,000 (32/16)	115	70	42	21	10/1	1			5	3	1	
ITF $50,000+H (32/16)	100	60	36	18	9/1	1			5	3	1	
ITF $50,000 (32/16)	80	48	29	15	8/1	1			5	3	1	
ITF $25,000+H (32/16)	60	36	22	11	6/1	1			2			
ITF $25,000 (32/16)	50	30	18	9	5/1	1			1			
ITF $15,000 (32/16)	25	15	9	5/1	0							
ITF $10,000 (32/16)	12	7	4	2	1/0							

Notes: +H indicates that Hospitality is provided. *Assumes undefeated Round Robin match record.
[a]W = Winner; F = Final; SF = Semifinal; QF = Quarterfinal; QLFR = Quali Loser Final Round.