Incentives to Quit in Men's Professional Tennis: An Empirical Test of Tournament Theory

Will Walker

Professor Peter Arcidiacono, Faculty Advisor
Professor Grace Kim, Seminar Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University
Durham, North Carolina
2018

Acknowledgements

I would like to thank Grace Kim, my seminar advisor, for her limitless patience, guidance, and support through the thesis-writing process. I would also like to thank Peter Arcidiacono, my advisor, for his invaluable guidance and feedback in shaping the empirical analysis of this paper. Finally, thank you to my peers Rahul Sharma, Grace Peterson, Nicholas Vega, and Rafal Rokosz for their continued feedback and support through the past year.

Abstract

This paper studies the influence of incentives on quitting behaviors in professional men's tennis tournaments and offers broader implications to pay structures in the labor market. Precedent literature established that prize incentives and skill heterogeneity can impact player effort exertion. Prize incentives include prize money and indirect financial rewards (ranking points). Players may also exert less effort when there is a significant difference in skill between the match favorite and the match underdog. Results warrant three important conclusions. First, prize incentives (particularly prize money) do influence a player's likelihood of quitting. Results on skill heterogeneity are less conclusive, though being the "match favorite" could reduce the odds of quitting. Finally, match underdogs and "unseeded" players may be especially susceptible to the influence of prize incentives when considering whether to quit.

JEL Classification: J41, J31, J32, J33, M12, M51, M52

Keywords: Tournament Theory, Incentives, Tennis, Quitting, Compensation, Sports Economics

Contents

1. Introduction

On July 4th, 2017, spectators at the Wimbledon Tennis Championships could not believe their eyes when, back-to-back, the opponents of Roger Federer and Novak Djokovic quit their matches shortly after they began, citing injury. Over the next two days of the famously-lucrative tournament, seven different players quit matches prematurely. Sports commentators, analysts – even casual spectators – began to question: did these professionals all suffer freak injuries on the court, or did they, when weighing the option of continuing to compete or bowing out and accepting a $50,000 guaranteed check, choose the latter? This paper addresses the resulting economic question: do incentives in men's professional tennis tournaments encourage players to quit?

The past 35 years has seen an explosion of interest in tournament theory as a proxy for measuring the effectiveness of incentive structures in competitive markets. At the core of the debate is a seminal economic question: does compensation based on competition make workers perform better? The answer to that question – regardless of its outcome – can change the way the economic community conceptualizes how an efficient organization should pay its employees.

In 1981, Lazear and Rosen made the landmark discovery that compensation schemes which pay according to an individual's rank in an organization, rather than individual output level, can induce the same efficient allocation of resources as a reward scheme based on individual output (Lazear & Rosen, 1981). Hundreds of subsequent extensions, both theoretical and empirical, tested Lazear & Rosen's hypothesis and extend its influence. Theoretical work has resulted in a more general hypothesis for tournament theory: competitor effort exertion increases with the spread between the winning and losing prize (Connelly, Tihanyi, Crook, &

Gangluff, 2014). This hypothesis has been empirically tested in fields ranging from academia, to criminal activity, to the labor market, and more (Connelly et al., 2014).

An important subset of economics has tested tournament theory through the vehicle of professional sports. Ronald G. Ehrenberg and Michael L. Bognanno's foundational paper, "Do Tournaments Have Incentive Effects?", found evidence that tournaments with higher prize offerings did lead to better scores from professional golfers (Ehrenberg & Bognanno, 1990). Subsequent literature, which spans activities like the NBA league (Grund, Höcker, Zimmermann, 2010), NASCAR racing (O'Roark, Wood, & Demblowski, 2012), and even tennis (Sunde, 2009; Ivankovic, 2007; Gilsdorf & Sukhatme, 2008; etc.), sought to empirically build upon these initial conclusions and extend their influence.

While, according to Connelly et al. (2014), major testable hypotheses of tournament theory have "enjoyed appreciable empirical support" (p. 20), tournament payouts must abide by certain conditions to incentivize optimal behaviors from contestants. In an addendum to his original work, Rosen offered a theoretical condition that is central to this thesis: when a tournament's guaranteed loser's payout is greater than the expected benefit of continued play, it is optimal to exert no effort – to quit (Rosen, 1986).

Tennis is uniquely situated for the study of incentive mechanisms in rank-order tournaments. Because information on relative player skill is available through regularly published rankings, it is possible to analyze how both prizes and difference in skill influence effort exertion in pairwise, sequential-elimination tournaments. Uwe Sunde (2003) discovered that both monetary incentives and differences in skill incentivize players to change their levels of effort in matches. Increasing the prize spread should increase effort from both competitors, but high skill heterogeneity should lower effort exertion from both competitors (Sunde, 2003).

Subsequent literature investigated the extent to which incentives versus skill heterogeneity could account for differences in effort exertion and match outcomes (Lallemand, Plasman, and Rycx, 2007; del Corral & Pireto-Rodriguez, 2010; Gilsdorf & Sukhatme, 2008).

This thesis shifts the focus of empirical study from optimal effort exertion to quitting behaviors. Just as incentives can influence how much effort a player exerts, can they influence whether he quits? If so, what exact effect do they have, and which players do they influence? This thesis empirically studies the role of incentives on quitting behaviors by creating a new dataset to examine matches on the men's Association of Tennis Professionals (ATP) tour from 2007 – 2017. Analyzing a combined dataset allows this thesis to make conclusions about the general role of incentives in influencing player quitting behaviors. Analyzing only the most lucrative tournaments on the tour – the "Grand Slams" – allows this thesis to perform robustness checks on the general results and gauge how different types of players respond to incentives to quit when these incentives are most exaggerated.

Undertaking this analysis may have important implications in the labor market. Of particular relevance is the tradeoff all workers in a firm must balance: the guaranteed base salary against the bonus that comes at the cost of effort exertion in the hope of advancement. If a worker's base salary is sufficiently high, why should he or she care at all about suffering the costs of entering the promotion or bonus "tournament" that promises increased cost of effort without promising any reward?

In the following sections, a background overview of the structure of the ATP tour is provided and precedent literature is integrated.

2. Background

2.1 ATP Tournament Layout and Incentive Structures[1]

The men's professional tennis tour is administered by The Association of Tennis Professionals (ATP) and the International Tennis Federation (ITF), which organize (as of the end of 2017) 68 yearly tournaments that range from 32 to 128 in draw size (ATP, 2017). In every professional tennis tournament (except the Davis Cup and the Olympic Games), players compete for both ranking points and prize money.[2] Better-ranked players qualify for more lucrative tournaments and may be "seeded" in a given tournament, which grants them a favorable draw, and by extension, a greater opportunity to accumulate both ranking points and prize money (Silverman & Seidel, 2011). If a player is ranked worse than the ranking "cutoff" for a given tournament, he must successfully pass through a qualifying tournament to gain entry. That being said, exceptions to this rule exist (Silverman & Seidel, 2011). These include being granted a "wild card," in which the tournament directors choose particular individuals for entry (usually home-country favorites or promising young players), or using a "protected ranking" to enter a tournament. Protected rankings allow injured players to enter a fixed number of tournaments using their pre-injury ranking to assist in attempts to return to the sport.

Professional tournaments on the ATP and ITF tours are sorted into 4 categories: "Grand Slam" tournaments, "ATP Masters 1,000" tournaments, "ATP World Tour 500" tournaments, and ATP World Tour 250" tournaments. Grand Slams are among the largest, most lucrative, and

---

[1] This section is inspired by a past Duke Honors Thesis written by Silverman & Seidel (2011) who provided a similar background in their work. They focused on how "intangible" incentives (such as home country advantage) impact player effort exertion.

[2] Figures 1 and 2 provide an example of prize money distribution and a visual representation of ranking point distribution across tournament classifications.

most prestigious tournaments, awarding 2,000 ranking points, millions of dollars in prize money, and elite recognition to their winners (ATP, 2017). Masters 1,000 tournaments comprise the next tier in terms of "status," awarding 1,000 ranking points to winners (ATP, 2017). ATP World Tour 500 and 250 tournaments are considered the lower-tiered tournaments, offering 500 and 250 ranking points to winners, respectively (ATP, 2017).

The incentive structure of a tournament is among its most important features. Prize incentives include both direct financial incentives (prize money) and indirect financial incentives (ranking points). As a reward for qualifying for a tournament (either through ranking alone or by surpassing a qualifying tournament), players receive a fixed sum of money, even if they lose their first match. As players advances to the next round(s), their marginal payouts increase.

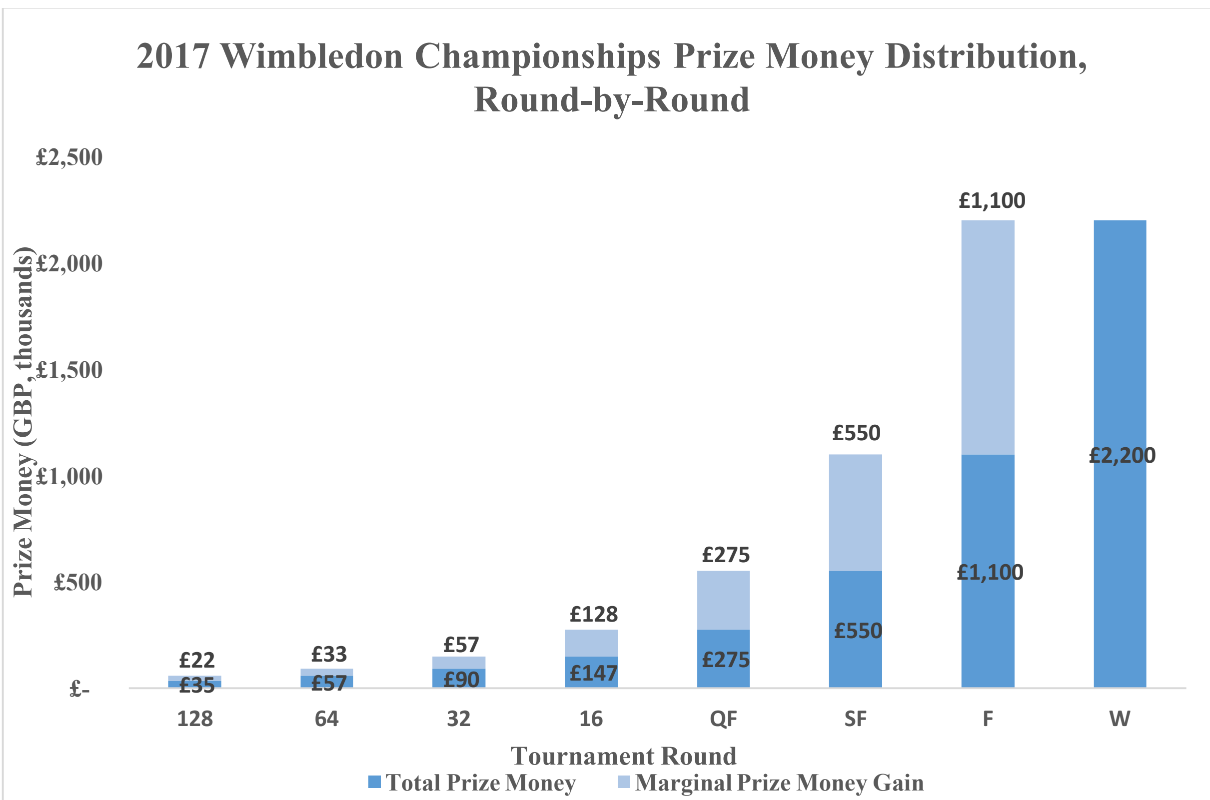

Figure 1: 2017 Wimbledon Championships Purse: Guaranteed Prize Money vs. Marginal Prize Money Gain (Source: The Wimbledon Championships)[3]

[3] Inspired by a similar graph created by Silverman & Seidel (2011).

Likewise, players receive an increasing number of ranking points for advancing to later rounds, though many tournaments do not guarantee a fixed point payout for first-round losers. While ranking points themselves are not monetary, a better ranking allows players to qualify for more lucrative tournaments and potentially have easier draws, increasing the likelihood of amassing prize money in the future.

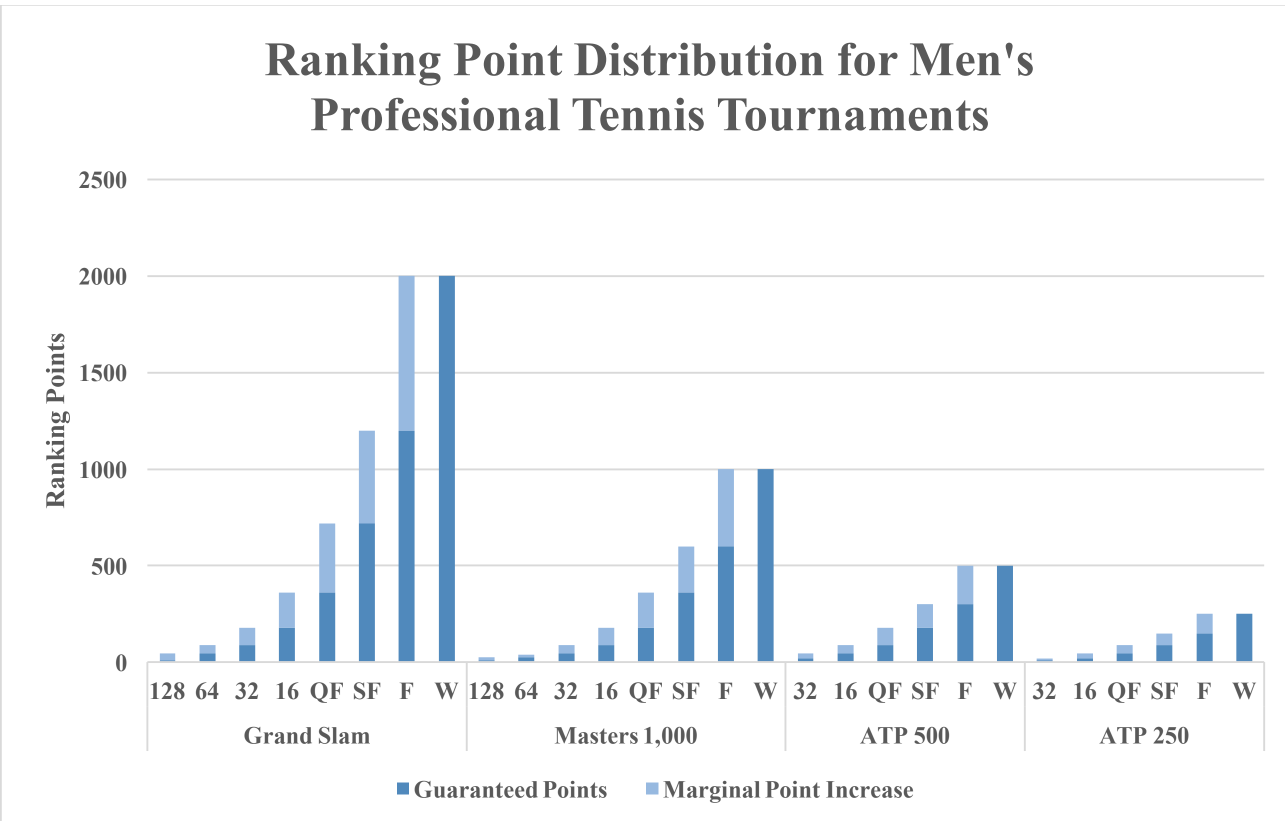

Figure 2: ATP and ITF Ranking Point Distributions by Classification, Guaranteed Points vs. Marginal Points Increase (Source: ATP World Tour)

2.2 Quitting Behaviors in ATP Tour Tournaments

At the start of any tournament, players have three potential courses of action: compete in the tournament and exert effort normally, enter into the tournament but quit during it, or pull out before competing. The ATP refers to instances when players quit as “retirements.” A player may retire during a match or in between rounds. **As such, the words "retire" and "quit" will be used interchangeably in this paper.**

There are a number of reasons why a player would choose either to quit or pull out before a professional tennis tournament. If a player becomes severely sick or injured prior to a tournament (e.g., breaking a bone during practice, being diagnosed with a serious illness, etc.) the physical exertion of competing in a tournament would likely compromise the player's health for the future, so he would be wise to pull out beforehand. If the injury or illness is not as serious (e.g., a muscle strain, the flu, etc.), a player may reasonably "give the tournament his best shot," but quit upon realizing his injury or illness is too severe to warrant continuing. Likewise, certain players may be more prone to said health concerns (older players, players with a past history of injury, and so on). Finally, players could quit tournaments for any number of significant personal reasons (family emergencies, etc.). These cases are, surely, perfectly reasonable. Nonetheless, the recent string of mid-match retirements at lucrative, high-profile events has caused onlookers to wonder if incentives may (at least partially) influence player quitting decisions. In a conversation with the New York Times (2018), Lukas Lacko, ranked 86th in the world at the time, noted:

> You work hard all season to get top 100, gluing together small results… [the first round loser's prize] is a reward you deserve because you got to top 100, and it's big money at these tournaments [a world-ranking within the top 100 generally guarantees entry into lucrative professional tournaments]. And if you suddenly, unluckily get injured, then you can't play and get paid where you were working hard for to play: the top of the hill. And if you don't play the match, you got zero. It was a big difference (paragraph 13).

Both the ATP and the ITF have made recent amendments to prize money rules regarding pulling out before tournaments. In 2017, the ATP tour instituted an experimental policy allowing players to pull out before 2 ATP tournaments per year and keep the first-round loser's prize money (ATP, 2017). In 2018, the four-member Grand Slam board announced it would allow injured players to pull out in the days before a Grand Slam and keep 50% of the first round loser's prize money, with the remaining purse distributed to the alternates who replace them (Waldstein, 2017).

It is too soon to tell if rule changes instituted by the ATP and ITF are effective in reducing quitting behavior on tour. As such, this work is concerned with whether status quo incentive structures of professional tennis tournaments influence players' likelihoods of quitting. That being said, future researchers may build upon this work by studying how the aforementioned rule changes impact both player quitting behaviors and the influence of incentives on said behaviors.

3. Literature Review

3.1 Theory

Tournaments as economic models were originally studied in Edward P. Lazear and Sherwin Rosen's foundational 1981 work, "Rank-Order Tournaments as Optimum Labor Contracts." The authors described two ubiquitous compensation models: "Piece rate" compensation and "rank-order" (or "prize") tournament compensation (Lazear & Rosen, 1981). The former evaluates an individual's performance against his or her *own* prior performance, and the latter evaluates the individual's performance against his or her competitors in a tournament format. As aforementioned, Lazear and Rosen argued that compensation based on a worker's relative position in the firm can produce the same efficient allocation of resources as a piece-rate compensation scheme provided that ability is known in advance and workers are risk neutral; in some cases, risk averse workers may even prefer being paid based on rank (Lazear & Rosen, 1981).

Since that seminal work, economists have made important theoretical expansions to rank-order tournament theory. Nalebuff and Stiglitz (1983) developed a number of extensions to the body of knowledge, among the most relevant of which are: (1) "the use of a contest as an incentive device can induce agents to abandon their natural risk aversion and adopt riskier, more profitable production techniques," (2) in contests with "low risk and, hence, small prizes, no symmetric pure strategy Nash Equilibrium will exist," and (3) when there are many competitors in the tournament, a penalty to the worst-ranked player will be superior to a prize to the best-ranked individual in motivating effort (p. 23). In 1994, Knoeber and Thurman used an empirical test to discover that changes in the level of prizes that leave prize differentials unchanged will not affect performance (Knoeber & Thurman, 1994). In other words: results suggested that only prize differential, not absolute prize level, matters for competitor incentives. Subsequent theoretical and empirical work tended to confirm the validity of tournament theory, at least for a time (Connelly et. al., 2014).[4]

That being said, certain work has outlined the limitations of tournament theory. In 2002, Hans K. Hvide (2002) found that Lazear-Rosen tournaments can bring about excessive risk and low levels of effort, a concept he coined the "risky-lazy trap" (paragraph 7). Certain management scholars have found that, while large prize spreads may motivate higher effort, a consideration of both economic factors (i.e. self-interest utility maximization) and behavioral factors (i.e., collaborative behavior) was needed to fully explain results (Henderson & Frederickson, 2001).

The key foundation for this thesis' theoretical framework is Rosen's seminal 1986 work, where he extended the rank-order tournament from the simplified two-player model to a sequential model in which players advance through rounds in a dynamic format. The results of his exercise found that, in certain conditions, quitting is the optimal decision (Rosen, 1986). He justified his conclusion by explaining that a player's incentive to exert effort boils down to a tradeoff between accepting the guaranteed loser's prize and continuing to compete for possible future compensation (Rosen, 1986). In cases when the guaranteed prize is greater than the expected benefit of continued participation, quitting is clearly optimal. This discussion is explored in more detail in the theoretical framework section.

[4] Economic work has also been undertaken to theoretically adapt tournament theory to the specific case of firm promotions. Emre Ekinci and Michael Waldman developed a new model that combined "classic" and "market-based" tournament theories to make predictions involving the relationship of bonus with job tenure, age, and more. They tested it empirically, and all predictions were confirmed (Ekinci & Waldman, 2015).

3.2 Empirical Research: Non-Tennis Activities

Particular empirical subjects of interest are professional golf, the NBA league, professional poker, NASCAR racing, and firm compensation policies. Ronald G. Ehrenberg and Michael Bognanno's 1990 work, "Do Tournaments Have Incentive Effects," is one of the earliest and best-known examples of viewing tournament theory through the lens of professional sports. Results from player performance on the 1987 European Professional Golf Association tour showed that increasing a tournament's total purse caused players to shoot better scores, especially in the tournament's final rounds (Ehrenberg & Bognanno, 1990).

More recent work has examined the potential for rank-order tournaments to induce risky behavior from their competitors, in line with theoretical concerns discussed earlier. An analysis of a year in the NBA league found that tournament competitors tended to increase their risk (the number of three-point shots attempted) if they were behind in a game, and that this behavior was only beneficial in increasing the probability of winning if the team was losing by many points (Grund et al., 2010). In professional poker tournaments, a study concluded that competitors were more likely to play risky bets if expected winnings were high or expected losses were low (Lee, 2004). Economic work in the realm of NASCAR racing discovered that a driver's relative ranking in "the Chase" (a points-based ranking system that determines top drivers to compete in a playoff) largely influenced the number of accidents in a given race. However, once qualifiers for the playoff had been determined, relative ranking was not found to be a strong predictor for risky driving (O'Roark et al., 2012).[5]

---

[5] Related papers have also sought to connect theoretical and empirical conclusions regarding tournament incentives to numerous scenarios in the professional world. In his 1989 work, Lazear concludes that "when workers' rewards are based on relative comparisons, salary compression reduces uncooperative behavior that is detrimental to the firm…Within the relevant groups, some wage compression is efficient" (Lazear, 1989). In 2006, Jed DeVaro estimates a structural

3.3 Empirical Research: Tennis

Precedent literature in the realm of tennis has been primarily concerned with analyzing how incentives impact effort exertion in matches. In the case of prize incentives alone, this task is straightforward: tournament theory would predict that higher prizes for advancing in the tournament increase effort exertion from players (Sunde, 2003). However, a problem arises when studying the role of incentives on effort exertion in tennis tournaments with uneven competitor skill levels: does skill heterogeneity incentivize players to exert less effort or do players of worse skill perform worse merely because of their lesser skill? The incentive effect of heterogeneity hypothesis suggests that, in instances of high skill difference between competitors, both the favorite and the underdog are incentivized to exert lower effort; the "plain capability effect" hypothesis suggests that underdogs, by nature of having lesser skill, perform worse than favorites (Sunde, 2003). Precedent literature has attempted to separate the incentive effect of heterogeneity from the pure capability effect and determine whether prizes or difference in player skill more strongly influence effort exertion.

Uwe Sunde pioneered this field of study. In analyzing the total prize money and marginal reward for advancing in the final two rounds of tournaments on the ATP tour from 1990-2002, he showed that the financial payout for advancing in a tournament –monetary incentives – mattered (Sunde, 2003). He subsequently made the novel discovery that, beyond the plain capability effect, both favorites and underdogs were incentivized to exert less effort in matches with high skill heterogeneity (Sunde, 2003). In 2009, he revisited the topic to create an empirical

---

tournament model treating performance, wage spread, and promotions as endogenous, and finds results consistent with the prediction of tournament theory that workers are motivated by larger spreads. That being said, Lazear (1989), Ehrenberg & Bognanno (1990) and Hvide (2002) also warn about applying tournament theory to *executive compensation*, as executives can sabotage one another or exhibit, "risky, lazy" behavior.

model describing the "incentive effects of heterogeneity" (Sunde, 2009). When combined with his work in 2003, it is the basis for the empirical analysis of this thesis.

Many precedent works have taken interest in analyzing whether player effort is influenced by skill heterogeneity or whether the capability effect alone explains differences in match outcomes. Thierry Lallemand, Robert Plasman, and Francois Rycx (2007) analyzed 502 matches on the Women's Tennis Association (WTA), and obtained results in line with the hypothesis that "the outcome of a match is more linked to players' abilities than to players' incentives to adjust effort according to success chances" (p. 3). Julio del Corral and Juan Pireto-Rodriguez tested whether the past performance of two players (using rankings as a proxy) could predict match outcomes. They used probit regression models to determine that, among all tested variables, ranking differential was the most significant in predicting who would win a tennis match, as the higher ranked player won the majority of the time (del Corral & Pireto-Rodriguez, 2010). However, Keith Gilsdorf and Vasant Sukhatme tested Rosen's sequential elimination tournament model in the context of men's professional tennis tournaments, and found that increasing prize money differentials between rounds in a tournament increased the chance that the better-ranked player would win (Gilsdorf & Sukhatme, 2008). Likewise, Ovaska & Sumell (2014), in using total tournament purse, tournament classification, and match round as incentive measurements, found that "incentives [did] matter: the higher ranked players [excelled] in more meaningful matches" (p. 38).

Other economists have taken more interest in the role of various types of incentives on effort exertion. Ivankovic (2007) tested how marginal pay distributions in professional tennis tournaments affected effort exertion, adding a "spread" variable that calculated a discounted present value of all potential matches each player could compete in during the tournament, multiplied by the winning probability of each subsequent match (pp. 89-90).[6] Ivankovic found inconclusive results. In some cases, player effort exertion was related to the tournament's overall purse instead of marginal payoff; in other cases, player effort was dependent on both marginal payoff *and* total purse (Ivankovic, 2007). Silverman & Seidel (2011) attempted to gauge the role of "intangible" incentives in effort exertion by comparing Grand Slam tennis matches with the Davis Cup, a World Cup-style team event that does not financially reward its winners. In including measurements of such incentives as home country pride/advantage, the relative "rank" of the player's country, and more, Silverman & Seidel found generally significant results, suggesting that non-monetary, intangible incentives did meaningfully increase competitors' effort exertions (Silverman & Seidel, 2011).

This paper contributes to the body of knowledge in three ways. First, it is among the first empirical papers to apply the economic logic of the incentive effect – both of prizes and of heterogeneity – to the study of quitting as opposed to optimal effort exertion. To the best of this author's knowledge, the only other economics paper studying quitting behaviors in tennis focused on how personal performance and income impacted whether players would quit the professional tour *altogether* (Geyer, 2010). Second, and by extension, it is among the first empirical papers to study how quitting behaviors are influenced by indirect measurements of financial incentives (ranking points). Finally, this thesis is among the first empirical papers to offer a discussion of how players of different skill categories respond differently to prize incentive effects when considering whether to quit.

[6] Due to time constraints of this thesis and limitations of the variable, empirical analysis accounting for the "spread" variable is not included in this thesis. Future work may consider replicating the analysis here but replacing the winning prize variable with the "spread" variable.

As a closing note, it is perhaps worth noting that tournament theory has been applied to study subjects in a wide variety of fields not mentioned in this literature review. These subjects include law (the race of judges to become Supreme Court Justices), psychology, finance, and even the meat industry (how contractors fight to provide chicken to Perdue and Tyson) (Connelly et al., 2014). The widespread applicability of tournament theory to separate fields across academia and the business world suggests that the findings of this paper may have relevance beyond firm promotion structures.

4. Theoretical Framework

4.1: Theoretical Foundation

Because the tennis tournament is the prototypical sequential, single-elimination tournament, Rosen's 1986 model remains the most appropriate from which to analyze the influence of incentives on quitting behaviors in professional tennis. Accordingly, many precedent papers in tennis use or draw influence from the model to conduct analysis (Sunde, 2003; Gilsdorf & Sukhatme, 2008; Ivankovic, 2007; Silverman & Seidel, 2011).

The single-elimination tournament starts with 2^N competitors and advances through N stages (Rosen, 1986). At every stage in the tournament, half of the competitors win and the other half lose. Each "stage" (for the purposes of this paper, "round") is a one-versus-one, or "pairwise" match where winners move to the next round and losers are eliminated. The overall winner's prize, W_1, is awarded to the tournament's winner, who has won *N* matches (Rosen, 1986). All players who lose with *s* rounds remaining in the tournament are awarded prize W_{s+1}, meaning the marginal prize money for advancement takes the form: $\Delta W_S = W_S - W_{S+1}$ (Rosen, 1986). The total reward for a player's performance in a given tournament is the sum of the marginal prize spreads accumulated, ΔW_S, up to the losing stage plus the guaranteed prize W_{N+1}.

In order for a player to advance in a tournament (and collect more prize money), he must exert effort to win the match at hand. Rosen denoted that it makes sense to view the player's effort decision as a probability function that depends upon the skill and effort of both players. The probability of winning a given match, $P_s(I, J)$, can be defined for players I and J as:

$$P_s(I,J) = \frac{\gamma_I h(x_{si})}{\gamma_I h(x_{si}) + \gamma_J h(x_{sj})} \quad (1)^7$$

[7] *h(x)* is increasing in *x* and *h(0)* ≥ 0.

Where x_{si} and x_{sj} "denote the intensity of effort expended by players i and j in a match where s stages remain to be played," and "γ_I and γ_J represent their abilities or natural talents for the game" (Rosen, 1986, p. 702). In essence, the equation describes how a player can increase his chance of winning through effort exertion, given a known value for his opponent's efforts and talents. This seems like a good fit for the case of a tennis tournament, in light of the discussion in the literature review regarding the the influence of skill heterogeneity on effort exertion. Public rankings of players on the ATP tour are good measurements of player skill, which could affect how a player views his probability of besting his opponent.

The player's decision regarding what level of effort to exert largely involves a tradeoff between the expected marginal benefit of winning the given match against the guaranteed prize and associated costs of exerting effort in continued competition (Silverman & Seidel, 2011). In the case of the tennis tournament, these costs are both real and metaphorical: a player should logically experience greater "wear-and-tear", soreness, and propensity for injury when competing in more matches, but should also increase real costs by increasing their lodging, transportation, coaching, and maintenance fees by staying at that tournament for an additional day(s).

Thus, the player must decide upon an effort level that maximizes his expected marginal benefit for winning the match and minimizes his costs of effort. Rosen defined V_s (I,J) as the match value to player I or J with s stages remaining in the tournament. The expected value of playing in later tournament rounds is denoted as (EV_{s-1}),[8] the probability of winning the match is denoted as (P_s), the match losing prize is denoted as (W_{s+1}), and the player's cost of effort is

[8] For clarification, "EV_{s-1}(I,J) is a weighted average over J of VS_{s-1}(I J), where the weights are probabilities that the player will confront an opponent of type J in the next stage" (Rosen, 1986, p. 703).

denoted as (c(x_s)), with c'(x) > 0, c''(x) ≥ 0, and c(0) = 0 (Rosen, 1986). The match value thus takes the form:

$$V_s(I,J) = \max\left[P_s(I,J)EV_{s-1}(I) + (1 - P_s(I,J))W_{s+1} - c(x_{si})\right] \quad (2)$$

4.2: Application – When is Quitting Optimal?

In using the foundational premises of this theoretical model to work out an exercise in which the decision to exert effort is *binary*, it is easy to see under which conditions quitting becomes the optimal response. Suppose for the moment that the contestant faces the following decision: (1) exert effort or (2) quit. In this isolated case, it is logical to normalize the guaranteed loser's payout (the "quitting" prize) to zero. The expected payoff for winning, then, may be represented by the following expression:

$$P_s(I,J)EV_{s-1}(I) - c(x_{si}) \quad (3)$$

where, as usual, $P_s(I,J)$ is the probability of the player in question winning the match, $EV_{s-1}(I)$ is the expected payout for continued participation (which includes both the immediate winner's prize and the expected participation in future match over a weighted average of different types of players), and $c(x_{si})$ is the cost of continued effort in the tournament. Thus, when:

$$P_s(I,J)EV_{s-1}(I) - c(x_{si}) > 0, \quad (4)$$

the player is better off exerting effort. Likewise, when:

$$P_s(I,J)EV_{s-1}(I) - c(x_{si}) < 0, \quad (5)$$

it is clearly optimal to exert zero effort – to quit.

A couple of important notes follow. First, the player's decision to quit depends upon three key factors: his probability of winning, his expected payout for winning, and his cost of effort. Second, Rosen structured his analysis to assume risk-neutrality (Rosen, 1986). This may not be the best categorization for professional tennis players; because the potential for lost income as a result of injury is a high-stakes proposition in professional tennis, it seems reasonable that (at least some) players could be risk-averse. To account for possible risk-aversion, Rosen concluded that tournaments must offer increasing marginal prize gains, particularly between the top and runner-up prizes (Rosen, 1986). Figures 1 and 2 show that ATP tournaments generally offer increasing marginal prizes, though the *rate* of marginal payout increases is not itself increasing. In theory, this could contribute to quitting tendencies, particularly among risk-averse players.

When the probability of winning, holding expected winner's payout constant, decreases, the likelihood of the player quitting increases and vice versa (the expression will potentially switch from positive to negative). This makes economic sense. As analyzed in Sunde's discussion of incentives in tennis, players of lesser skill have smaller probabilities of realizing gains from the winner's prize, and are subsequently motivated to exert less effort. In extreme cases, such as injury or when the disparity between the skill of the player and his opponent is sufficiently large, the probability of winning may become so low that quitting the match is an optimal decision.

A similar discussion holds for expected payout: when a player's expected future winning prize (holding the probability of winning the match at hand constant), discounted over the course of the tournament, decreases, the player's likelihood of quitting increases and vice versa. Here, it is important to note that the expression for expected winning prize considers the winning prize for the match at hand and the prizes for all subsequent matches that the player could potentially compete in, multiplied by the probability that he wins each of those successive matches. Of course, the immediate winner's prize is most important to the player, and one would logically expect that payouts for winning the match at hand would have the most influence on player effort exertion. But, while an injured or comparatively low-ranked player may have a legitimate shot at winning the match at hand, he may have disproportionately low chances of winning future matches in the tournament. This is important if the player is keen to view the tournament as an entire entity rather than a collection of individual matches.

Finally, holding both probability of winning the given match and expected future winnings constant, increased cost of effort increases the likelihood that a player will quit, and vice versa. Economic logic and common sense affirm that this is true. Consider, again, the case of an injury. Regardless of whether a player is the favorite to win the match, the next couple of rounds, or even the whole tournament, injury increases the cost of effort by decreasing the chance that the player will be healthy to compete in subsequent rounds or subsequent tournaments. If a player's injury is sufficiently severe (e.g., an ankle sprain, a torn ligament, etc.), it is surely optimal to quit now and avoid risk of even greater injury that could compromise winnings well into the future. Another component of cost of effort, real costs like hotel, coaching, and maintenance fees, are also important. These costs disproportionately impact "journeymen" players who do not have the financial resources of the game's elite. For them, taking on continued real costs in a tournament may only be justified if expected winnings are sufficiently large.

The final discussion reveals an important limitation of Rosen's model. In his conception (1986), cost of effort was assumed "identical for all players" (p. 703). In professional tennis tournaments, this is highly unlikely to be the case. In particular, costs among players in a tennis tournament may differ in two important ways: in the absolute magnitude of costs and in how much those costs matter to players' bottom lines. First, worse-ranked players may experience higher "wear and tear" cost than better-ranked opponents because they typically play better players earlier in tournaments, and could exert more effort per match.[9] Likewise, unseeded players are not awarded "byes," and may have to play more matches to advance to the same round as their seeded counterparts. Second, as discussed earlier, worse-ranked players typically have lower career prize money ("journeymen" players) and, as a result, could place more value than wealthier competitors on guaranteed payouts, as these payouts help players cover vital fixed costs like hotel, coaching, and equipment fees. The higher relative "investment cost" of less-wealthy players in a given tournament is likely a vital consideration in the decision to quit during tournaments rather than pull out beforehand. The following section discusses the problems of constructing a variable to account for costs in further detail, as well as how that may impact empirical analysis.

[9] The structure of a tennis tournament ensures this is the case. Where better-ranked players are frequently "seeded" and play worse-ranked players in early rounds, unseeded players typically draw "seeded" players early on.

5. Data and Variables

5.1 Initial Remarks

A unique dataset was created for this thesis. Publicly available records on professional match scores, tournament prize-money levels, and individual match performance (i.e. winner/error ratios, first serve percentages, etc.) are available on www.atpworldtour.com (ATP). Each data point, denoted by a row in the dataset, represents an individual player in a tournament match, and includes information about the match's court surface, year, its given tournament, the difference in ranking between the player and his opponent, the prize money associated with the given match, and more.[10] The dependent variable of interest in this paper, whether or not a given player quits (retires), is noted in scores that the ATP reports.

While similar datasets have been used in tennis papers, this particular dataset is unique and has not been used in academic analysis before. It is the belief of this author that because of its comparatively robust size and collection of more recent observations, this dataset will allow for more accurate, nuanced analysis. For example, Ivankovic's dataset included 169 professional tennis tournaments from the 1992 and 1993 seasons (Ivankovic, 2007). Sunde used data from 156 men's tennis tournaments from 1990-2002, and included only matches from the last 2 rounds of tournaments (Sunde, 2003). Gilsdorf and Sukhatme analyzed a dataset of 68 tournaments, encompassing 2,632 matches during the 2001 season of the ATP tour (Gilsdorf & Sukhatme, 2008). This thesis's dataset includes measurements of 29,547 player-specific

[10] Specific variable descriptions for each data point are explained in greater detail in section 5.3. below.

observations in matches (in the combined dataset), drawn from 391 tournaments in the 2007-2017 seasons.[11]

5.2 Dataset Source, Scraping, and Cleaning Techniques

As aforementioned, all relevant data points are drawn from publicly available data on www.atpworldtour.com. The data cleaning process is summarized in Figure 4 below.

Figure 4: Dataset Construction and Cleaning

Data Cleaning Process	**Observations Remaining**
Construction of total initial sample of select tournaments 2007-2017 (by match)	19,089
Reorganized sample of select tournaments 2007-2017 (by player)	38,178
Drop tournaments without recorded prize money distributions	37,614
Drop matches with unranked players	36,536
[Stata] Drop all round 7 observations (perfect prediction because no players quit)	36,406
[Stata] Drop all players who have never quit (perfect prediction) or who have played fewer than 50 matches (Stata calculation problems)	29,547
Combined dataset size	**29,547**
[Stata] Drop all non-Grand Slam tournaments and results from rounds 6 and 7 (perfect prediction because no players quit)	6,185
Grand Slam dataset size	**6,185**

A more detailed description of the specific steps taken to construct the dataset is included in addendum 1 of the appendix.

5.3 Construction of Dataset Variables

A summary of all currently-utilized variables can be found in Table 1 of the appendix. Variables were constructed to capture the core tenets of the theoretical model and to account for

[11] The observation number is odd because, as noted in the data cleaning discussion, players of certain types were dropped to enable successful calculation of the model.

issues salient to the specific case of player quitting behavior: skill heterogeneity, winning and losing prizes (the incentive effect), cost of effort and investment cost, and propensity for injury.

To account for skill heterogeneity, variables *BetterRank, RankDiff,* and *RankDiff***BetterRank* were constructed. *BetterRank* is a binomial variable denoting whether the player has a better prior year-end rank than his opponent (1 if better-ranked), *RankDiff* is the absolute value of the difference in ATP ranking between the favorite and the underdog, and *RankDiff***BetterRank* is the interaction between the two. Incorporating variables that account for skill heterogeneity allows this paper to delineate how both incentives and the probability of winning a match influence quitting behaviors. Before continuing, is is important to note that there are some potential problems with using the above heterogeneity variables. First, they pull the player's year-end rank in the year prior to a given match, subsequently ignoring any improvements a player made in the same year as that match. For example, this technique may not accurately predict quitting behaviors in 2017 for players who had significant breakthrough results in 2017 (e.g., winning or advancing to a final round in a prestigious tournament in 2017).

To account for prize incentives, the following variables were constructed: *WPM, WPo, GPM,* and *GPo.* If M_i is the payout a player is guaranteed for making it to round *i* of a tournament, his marginal payout for advancement is simply M_{i+1} - $M_{i.}$ In this model, *GPM* is M_i and *WPM* is M_{i+1}. Both values are divided by $1,000 to facilitate interpretation of the empirical model. *GPo* and *WPo* are the ranking point equivalents. Unlike precedent literature, this thesis deconstructs marginal prize into its two components because when a player is considering whether to quit, he is likely not only concerned with the prize spread, but also the payout he is guaranteed for making it to that given stage. *WPM* and *WPo,* for interpretive purposes, function similarly to the marginal prize variable. The incentive effect would predict that, as guaranteed payout increases and winning payout decreases, a player is more likely to quit and vice versa.

Measurements for important theoretical and practical concepts specific to quitting in professional tennis are important to include in the regression to avoid omitted variable bias, but are difficult to measure. Key examples are player cost of effort (as discussed in the theoretical framework section) and player propensity for injury. Both are difficult to systemize. Most players have different coaches, use different equipment, travel to different tournaments, and hail from different regions. Computing a function that captures the sheer breadth of influences on player cost may not be possible. A measurement of player injury would also be important to include in the regression, but is similarly difficult to calculate. A suitable proxy would be to count the number of times a player pulled out before tournaments in the 2-3 months leading up to a given match, as in these cases the player is sufficiently injured to completely forego participation and payouts. However, instances of players pulling out before tournaments are difficult to collect in aggregate because all players who do so are replaced with alternates and removed from draw records. Because of said difficulties, variables for cost of effort and injury were not constructed for analysis. Future research may improve upon the model developed in this thesis by constructing appropriate methods to estimate these factors. To deal with potential omitted variable bias, this work instead constructs player, match, and tournament fixed effects estimators. These are discussed in the empirical specification section.

5.4 Data Discussion, Correlational Relationships, and Limitations

A histogram of the data is shown in Figure 5 below.

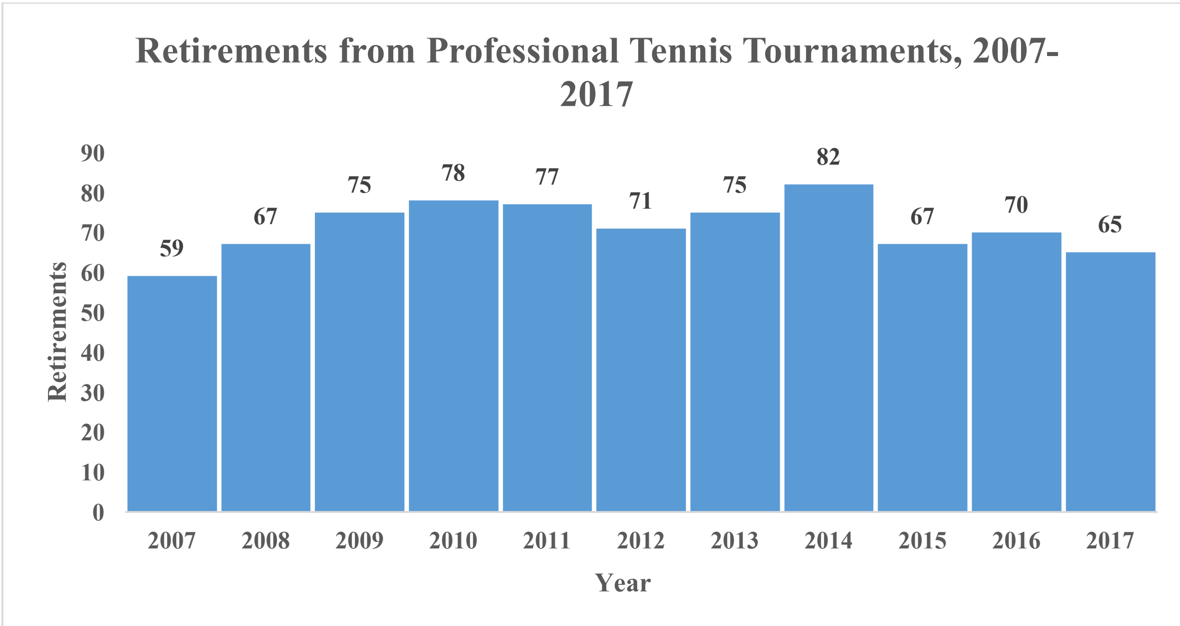

Figure 5: Histogram of yearly player quitting occurrences, 2007-2017

The distribution of quitting behaviors year-over-year is relatively uniform, suggesting that there is not an immediately-discernable trend over time and that the data may not fit a normal distribution well. This is further discussed in the Empirical Specification section. Descriptive statistics of the combined and Grand Slam datasets are included in tables 3 and 4 of the appendix. The statistics generally show that the standard deviations of the prize incentive variables are greater than their respective means. This is consistent with the notion that marginal prize increases significantly each round, contributing to incentive values that are highly spread about the mean. Mean values of player skill classifications, *BetterRank* and *Seeded*, are both greater than 0.5, suggesting that more matches involve seeded players and favorites. This makes sense; match favorites and tournament seeds are likely better than "average" professional tennis players, causing them to win and compete in more matches. Finally, mean values on *Retire* suggest that players quit 2.3% of matches in the general dataset and 3.6% of matches in the

Grand Slams dataset, consistent with the notion that players quit more in the most lucrative tournaments.

Chi-squared tests were run to understand correlational trends of quitting behaviors and how they vary across different tournament payouts and player skill types. Results help inform whether incentives and skill heterogeneity are relevant for the multivariate analysis undertaken later in the paper. With regard to tournament-level analysis, it is simplest to segment tournaments by total prize money. To divide tournament purse roughly along classification lines,[12] 5 total purse categories were created. The results of the chi-squared test are printed below. The test confirms that tournaments with the highest total prize money ($10M-$25M) have a higher percentage of matches quit than other groups, which do no appear to have large differences in rates of quitting. The associated p-value indicates that the relationship among these variables is significant at the 5% level.

Tournament Prize Money Range (USD, Thousands)	**Matches Quit (%)**
$0-$750	1.86%
$750-$2,500	2.02%
$2,500-$5,000	1.97 %
$5,000-$10,000	2.03%
$10,000-$25,000	**2.51%**

Pearson $\text{Chi}^2 = 10.3725$, Pr = 0.035

Figure 6: Chi-squared analysis results, total tournament purse segmentation

The next logical step is to understand what kinds of players quit, based upon skill heterogeneity. This can be broken down further into three factors: (1) player quitting differences *within matches*, (2) player quitting differences *within tournaments*, and (3) whether players of

[12] As aforementioned, classifications include: ATP 250, ATP 500, Masters 1,000, and Grand Slam tournaments.

worse year-end ranking are more likely to quit *in general*. Consider first skill differences within matches. For illustrative purposes, assume there is one simple distinction between players: whether they are the match favorite or the match underdog.

Favorite or Underdog?	Matches Quit (%)
Favorite	1.97%
Underdog	**2.25%**

Pearson $Chi^2 = 3.5873$, $Pr = 0.058$

Figure 7: Chi-squared analysis results, match favorite vs. underdog segmentation

This is obviously a simplifying assumption, but results suggest that match underdogs quit more frequently than match favorites. The chi-squared test results and the associated p-value suggest that the relationship between these variables is significant at the 10% level.

To understand how player quitting rates differ within tournaments (when accounting for skill heterogeneity), it is logical to compare "seeded" versus "unseeded" players. Seeded players are generally in the top-quarter (in terms of ranking) of all players in a tournament, with the remaining field filled by unseeded players. The seeding system acts as a proxy for measuring a player's "relative rank" in the tournament; seeded players would naturally expect to advance further in a tournament than unseeded players. Results of the chi-squared test and the associated p-value confirm that unseeded players do quit more often than seeded players do, and that the relationship among these variables is statistically significant.

Unseeded?	Matches Quit (%)
Yes	**2.56%**
No	1.73%

Pearson $Chi^2 = 31.1837$, $Pr = 0.000$

Figure 8: Chi-squared analysis results, unseeded vs. seeded segmentation

Likewise, segmenting the data by ranges of player rankings allows for analysis to determine which players are more likely to quit in general, regardless of tournament type. Five categories were created to gauge a player's skill, and the resulting chi-squared output is below.

Ranking Range	Matches Quit (%)
1-25	1.71%
25-50	2.25%
50-100	**2.70%**
100-200	1.92%
>200	0.97%

Pearson $Chi^2 = 40.177$, $Pr = 0.000$

Figure 9: Chi-squared analysis results, year-end rankings

The results suggest an interesting conclusion: the worst-ranked players who consistently gain entry to ATP tournaments (ranking range 50-100), have higher quitting rates than any other group, at 2.70%. Against expectations, players in category 5 had the lowest quitting rates, at 0.97%. This is likely due to the comparatively small number of observations in the group, combined with the fact that, for many of these players, qualifying for an ATP tournament is an exceedingly rare accomplishment to be savored. The chi-squared test results and associated p-value confirm that the relationship among these variables is statistically significant.

Results of correlational analyses reveal what types of players are more likely to quit and what types of tournaments and matches they are more likely to quit in. Generally, the most lucrative tournaments feature the highest quitting rates, suggesting that the incentive effect matters. Difference in player skill, whether measured within a match, within a field of competitors in a tournament, or within the general field of ATP players during the whole year, also results in different quitting rates, with players of worse relative skill quitting more than players of better relative skill. This suggests that skill heterogeneity also influences player quitting behaviors.

These analyses are too rudimentary to assign causal power, but do suggest that including these distinctions in the empirical model is necessary and could lead to interesting results.

Limitations of the data set relate to the nebulous nature of quitting in tennis tournaments, "messy" observations of tennis competitions, and limitations to the data collection process. Foremost, players may decide to quit for any number of reasons, including: injury sustained before the tournament, injury sustained during the match, illness, lack of fitness (cramping), general medical emergencies, personal/family obligations, and more. These factors are not identified in the data. This, combined with the fact that a player's decision to quit is an extreme case of an innately human behavior that could well be impulsive, means that constructing a systematic model to explain this behavior is difficult. Finally, qualitative, exogenous influences may skew the data and can be difficult to systemize. These include such factors as tournament "prestige," coaching quality, and extreme weather conditions (e.g., heat, coldness, humidity, etc.).

6. Empirical Specification and Analysis

6.1: General Remarks and Interpretation

This thesis utilizes a logistic specification with player, match, and tournament fixed effects estimators. While either a probit or logistic regression would achieve similar empirical results, the logistic specification was chosen due to the distribution of this thesis's dataset. As shown in Figure 5 and discussed in the data section, a histogram of quitting behaviors among men's professional tennis players, year-over-year, is relatively uniform, suggesting that the data would not fit a normal distribution well. The probit specification is based upon a cumulative standard normal distribution, whereas the logistic specification is based upon a cumulative standard logistic distribution. Because the logistic distribution captures greater kurtosis than the normal distribution does, a logistic regression would likely describe quitting behaviors in this dataset better than the probit specification would.

Because of the aforementioned difficulty of constructing adequate variables to account for player cost of effort and injury, fixed effects estimators are added to the logistic specification to account for unobserved player-to-player, match-to-match, and tournament-to-tournament differences. Player fixed effects are estimated by including an encoded string of all player names in the regression. They are likely the most important of the fixed effects estimators, because they control for minute differences among players, including individual cost of effort, propensity for injury, whether a player is more likely to "game the system," and so on. Match and tournament fixed effect estimators are also important to include because they account for any differences specific to a particular match or tournament that might contribute to quitting behaviors. Tournament and match fixed effects estimators are tournament draw size, court surface, round, and year of tournament.

The logistic specification generates odds ratio coefficients that are more easily-interpreted than typical logit outputs. Odds ratio coefficients denote how much more or less likely a dependent variable is to occur with a 1-unit change in the identified independent variable, holding other factors constant. Any coefficient value above 1 indicates that the dependent variable is more likely to occur when increasing the independent variable; any coefficient value below 1 indicates the dependent variable is less likely to occur when increasing the independent variable. Values close to 1 indicate that the independent variable has little influence on the dependent variable; values close to 0 or much greater than one indicate significant influence. For example, a coefficient value of 1.25 on the *GPM* variable (in thousands of dollars) would mean that, for every $1,000 increase in guaranteed prize money, a player is 25% more likely to quit, when controlling for all other factors.

Hypotheses and expected directionality on all terms included in the regression can be found in Table 2 of the appendix, and are summarized here. One would expect coefficients on guaranteed prize variables (*GPM* and *GPo*) to be greater than one and coefficients on winning prize variables (*WPM* and *WPo*) to be less than one; increasing guaranteed prize should increase likelihood of quitting and increasing winning prize should decrease likelihood of quitting. Coefficients on *Seeded* and *BetterRank* should be less than one; favorites should be less likely to quit than underdogs. The coefficient on *RankDiff* is expected to be greater than one; matches with greater skill heterogeneity should result in a higher likelihood of one of the players quitting. The coefficient on the interaction term of *RankDiff**BetterRank* should be less than one; underdogs, rather than favorites, should be more likely to quit in instances of high skill heterogeneity. Coefficients on the interactions of the guaranteed prize variables (*GPM* and *GPo*) with *BetterRank* and *Seeded* variables are expected to be less than one because, due to the higher "value" of prize money to underdogs and unseeded players discussed earlier, favorites are less likely to be influenced by incentives than underdogs are when considering whether to quit. For similar reasons, the interactions of winning prize variables (*WPM* and *WPo*) with the *BetterRank* and *Seeded* variables are expected to be greater than one.

6.2: Empirical Models

The base empirical model combines two specifications developed by Sunde (2003 & 2009) to evaluate the impact of both prize incentives (money and ranking points) and skill heterogeneity on quitting behaviors:

$$ln\frac{P(RET_{IMJ}=1)}{1-P(RET_{IMJ}=1)} = \beta_0 + \beta_1 PRIZE_{IMJ} + \beta_2 FAV_{IMJ} + \beta_3 HET_{IMJ} + \beta_4 FAV_{IMJ} * HET_{IMJ} +$$

$$\beta_5 X_I + \beta_6 Y_M + \beta_7 Z_J + \varepsilon_{IMJ}$$

where the dependent variable, *RET,* is a binomial measurement of whether player *I* retires/quits (1 if the player quits) in match *M* of tournament *J*, $PRIZE_{IMJ}$ is a vector of variables measuring the effects of prize incentives on quitting behaviors for player *I* in match *M* of tournament *J*, FAV_{IMJ} is a binary variable for whether player *I* is the underdog or favorite of match *M* (1 if the player is the favorite), HET_{IMJ} is a variable measuring the effect of skill heterogeneity on quitting behaviors, X_I is a dummy variable controlling for player fixed effects, Y_M is a vector of dummy variables controlling for match fixed effects, and Z_J is a vector of dummy variables controlling for tournament fixed effects.

In the base model, $PRIZE_{IMJ}$is accounted for by the variables *GPM, WPM, GPo* and *WPo*, which represent guaranteed payout and winning payout (both financially and in ranking points), respectively. FAV_{IMJ} is accounted for by the variable *BetterRank,* HET_{IMJ} is accounted for by the variable *RankDiff,* and the interaction term is accounted for by *RankDiff*BetterRank.* Player fixed effects are accounted for by including an encoded string of player names in the regression (*PCode)*. Match and tournament fixed effects are accounted for by including the round of the given match (*Round)*, an encoded string for the court surface of the match/tournament (*SCode)*, the year of the match/tournament (*Year),* and the tournament draw size (*DrawSize)*. Thus, the base model, denoted Model I, is represented by:

$$ln \frac{P(RET_{IMJ}=1)}{1-P(RET_{IMJ}=1)} = \beta_0 + \beta_1 BetterRank_{IMJ} + \beta_2 RankDiff_{IMJ} + \beta_3 RankDiff_{IMJ} * BetterRank_{IMJ} + \beta_4 GPM_{IMJ} + \beta_5 WPM_{IMJ} + \beta_6 GPo_{IMJ} + \beta_7 WPo_{IMJ} + \beta_8 PCode_I + \beta_9 Round_M + \beta_{10} SCode_J + \beta_{11} Year_J + \beta_{12} DrawSize_J + \varepsilon_{IMJ}$$

The model is extended to capture different measurements of player performance, allowing for study of how different types of players, based on skill, respond to changes in incentives when considering whether to quit. The first extension contrasts how match underdogs and match favorites respond to equivalent changes in financial and ranking point incentives.[13] In doing so, the model focuses on the immediate, short term quitting decision players make when they compare themselves to their direct opponent. The *BetterRank* variable is subsequently interacted with the four incentive variables. No other variables change. This specification, Model II, is:

$$ln \frac{P(RET_{IMJ}=1)}{1-P(RET_{IMJ}=1)} = \beta_0 + \beta_1 BetterRank_{IMJ} + \beta_2 RankDiff_{IMJ} + \beta_3 RankDiff_{IMJ} * BetterRank_{IMJ} + \beta_4 GPM_{IMJ} + \beta_5 WPM_{IMJ} + \beta_6 GPo_{IMJ} + \beta_7 WPo_{IMJ} + \beta_8 BetterRank_{IMJ} * GPM_{IMJ} + \beta_9 BetterRank_{IMJ} * WPM_{IMJ} + \beta_{10} BetterRank_{IMJ} * GPo_{IMJ} + \beta_{11} BetterRank_{IMJ} * WPo_{IMJ} + \beta_{12} PCode_J + \beta_{13} Round_M + \beta_{14} SCode_J + \beta_{15} Year_J + \beta_{16} DrawSize_J + \varepsilon_{IMJ}$$

The final extension of the basic model contrasts how unseeded and seeded players within a tournament respond to equivalent changes in financial and ranking point incentives. In doing so,

[13] The correlational relationship between match favorite vs. underdog and quitting behaviors is discussed in section 5.4.

it studies the more general quitting decision players make when they compare themselves against *the field* rather than against their immediate opponent (players know if they will be seeded before the tournament starts). The *Seeded* variable is interacted with the four incentive variables. No other variables change. This specification, Model III, is denoted as:

$$ln\frac{P(RET_{IMJ}=1)}{1-P(RET_{IMJ}=1)}=\beta_0+\beta_1 BetterRank_{IMJ}+\beta_2 RankDiff_{IMJ}+\ \beta_3 RankDiff_{IMJ}* BetterRank_{IMJ}+\ \beta_4 GPM_{IMJ}+\beta_5 WPM_{IMJ}+\beta_6 GPo_{IMJ}+\beta_7 WPo_{IMJ}+\ \beta_8 Seeded_{IMJ}* GPM_{IMJ}+\beta_9 Seeded_{IMJ}*WPM_{IMJ}+\ \beta_{10} Seeded_{IMJ}*GPo_{IMJ}+\ \beta_{11} Seeded_{IMJ}*WPo_{IMJ}+ \beta_{12} PCode_I+\ \beta_{13} Round_M+\ \beta_{14} SCode_J+\ \beta_{15} Year_J+\ \beta_{16} DrawSize_J+\varepsilon_{IMJ}$$

Seeded and *BetterRank* interaction effects are included in separate regressions because there is collinearity between the two variables, which could distort the results.[14] While they capture slightly different ways in which players differ in skill (against the field versus against an individual opponent), it is a natural trend that better-ranked players are seeded and that the worse-ranked players are not.

The models are applied to two different data segmentations: the combined dataset and a separate dataset containing only Grand Slams. The combined dataset allows this thesis to make general conclusions regarding how incentives impact players' decisions to quit. The Grand Slam dataset is helpful because it isolates tournaments with the most exaggerated incentive effects, making it particularly conducive to studying how player quitting behaviors differ when incentives are the highest. For reasons explained in the results section, it also supplements analysis regarding how indirect financial incentives (ranking points) influence quitting behaviors.

[14] See Table 9 for a correlation matrix of independent variables included in the regression.

7. Results and Discussion

Results of the regressions are included in Tables 5 and 6 of the appendix. Table 5 presents results for the combined dataset and Table 6 presents results for the isolated Grand Slam data. In both tables, the first column includes results from Model I without fixed effects estimators,[15] the second includes results from Model I with fixed effects estimators, the third includes results for Model II with fixed effects estimators, and the fourth includes results for Model III with fixed effects estimators.

7.1: Combined Dataset Results

7.1.1 Model I: Base Model

The first column of Table 5 is included to provide empirical evidence that including the fixed effects estimators improves the fit of the model in addition to making economic sense. Without fixed effects estimators, the model fits the data poorly. While its chi-squared statistic of 31.41 (with 7 degrees of freedom and a p-value of 0.000) indicates that there is a statistically significant relationship between the independent variables and the dependent variable, its Pseudo R^2 value is slightly less than 0.005, which seems extremely low even in the context of quitting. Including the fixed effects estimators increases the Pseudo R^2 value to 0.055 and the chi-squared statistic to 360.36 (with 182 degrees of freedom and a p-value of 0.000). This is more in line with the expectations of this paper. The increased chi-squared statistic demonstrates an increase in statistical significance of the model and the Pseudo R^2 value seems much more reasonable. Because quitting is a human decision that is often spontaneous and difficult to predict, this work

[15] Model I without fixed effects estimators includes more observations than Model I with fixed effects because, in controlling for player fixed effects, the model drops all players who have never quit a tennis match due to perfect prediction of player behavior.

does not expect Pseudo R^2 values to be exceedingly high. Accordingly, the following discussion only interprets results from Model I with fixed effects estimators.

Just as precedent literature has found that incentives influence effort exertion, results of the base model generally confirm the hypothesis that financial incentives influence quitting behaviors. Coefficients on direct financial incentive variables, *GPM* and *WPM,* are statistically significant at the 10% level with directionality consistent with the expectations of this paper. *GPM's* coefficient, 1.018, suggests that players are 1.8% more likely to quit matches when their guaranteed loser's payout increases by $1,000, controlling for other factors. *WPM's* coefficient, .989, suggests that players are 1.1% less likely to quit when their total prize for winning increases by $1,000. These effects may not seem particularly powerful without context, but when considering that guaranteed and winning prizes in the first round of tournaments alone can range from $2,000-$50,000 and $5,000-$80,000, respectively, it is clear that financial incentives can play a significant role in whether a player decides to quit.

Results for ranking point prizes – which may lead to higher future earnings but are not themselves financial incentives – behaved against the expectations of this paper. Results on *GPo* were statistically insignificant with opposite directionality of this paper's expectations. Results on *WPo,* significant at the 10% level, also had opposite directionality of expectations. Confounding factors unique to the ranking point variables could account for the unexpected results. For example, note Figure 10 below. ATP 250 and ATP 500 classification tournaments do not guarantee *any* ranking points for first-round losers; players must win their first-round matches to collect ranking points.

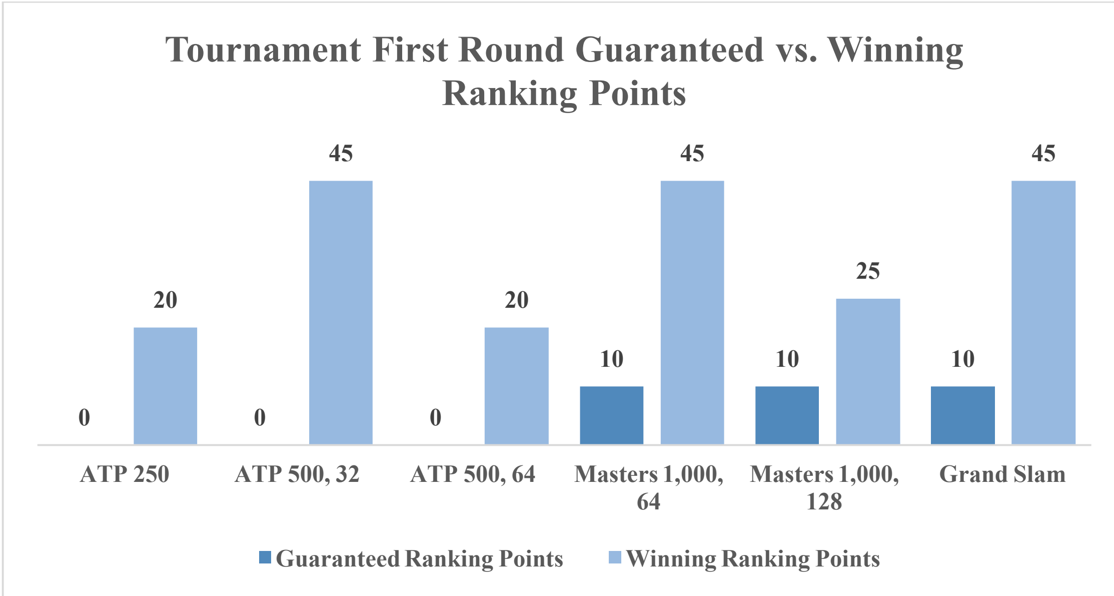

Figure 10: First Round Guaranteed and Winning Ranking Points by Tournament Classification and Draw Size (Source: ATP Tour)

As a result, the effects of financial incentives and ranking point incentives are confounded; while players may well quit during the first round of ATP 250 and ATP 500 tournaments to receive their guaranteed first-round loser's check, in doing so they forfeit all potential ranking point gains from the tournament. So, the influence of ranking point incentives on quitting could be "masked" by the fact that, when players consider whether to quit in these tournaments, they likely weigh guaranteed money and the lack of guaranteed ranking points against one another. In order to account for this problem, it is necessary to narrow the analysis to tournaments in which first-round losers are guaranteed ranking points in addition to prize money. As seen in Figure 10, the Grand Slam tournaments are such a category. Another benefit of running the model on the isolated Grand Slam dataset, then, is that it helps parse out the influence of ranking point incentives when they follow a similar distribution to financial incentives.

Results on variables measuring skill heterogeneity are likewise statistically insignificant and inconclusive. The variable *BetterRank* did exhibit expected directionality, suggesting that being the "favorite" may reduce the odds that a player will quit. However, coefficients on variables *RankDiff* and the interaction *BetterRank*RankDiff* are extremely close to 1. Such an outcome leads to two hypotheses. First, while the distinction of being the favorite may influence quitting likelihood, the *magnitude of skill difference* may not. Second, the relationship between quitting decisions and skill heterogeneity may be nonlinear, which could distort results. Consistent with precedent literature (Silverman & Seidel, 2011), all models (on both datasets) were run with the *RankDiff* variable squared ($RankDiff^2$). For reference, results of these regressions are included in tables 7 and 8 of the appendix. In all cases, coefficients remain statistically insignificant and close to one.

If one is inclined to believe the first hypothesis, results could contribute to the debate in precedent literature regarding whether prize incentives or skill heterogeneity impact effort exertion more; at least in the context of quitting, prizes may matter more. That being said, results could also indicate that the relationship between the two variables is neither linear nor quadratic. Instead, the influence of equivalent magnitudes of skill heterogeneity could be inherently varied based on player ranks. While it is completely reasonable to suggest that there is a significant "skill gap" between the world's number 1 ranked player and the world's number 50 ranked player, there is likely a much smaller "skill gap" between a player ranked 201 and another ranked 250. Future research may attempt to measure this hypothesized variation.

7.1.2 Model II: Base Model with Match Underdog Interactions

Interacting the *BetterRank* variable with the incentive variables allows this thesis to account for the possibility that the role of prize incentives on quitting behaviors is also influenced by whether the player is the match favorite or the match underdog.[16] The Pseudo R^2 value of 0.056 and the chi-squared statistic of 368.87 (with 186 degrees of freedom and a p-value of 0.000) suggest that Model II and Model I have similar statistical significance and goodness-of-fit.

In the interacted model, coefficients on incentive variables now represent how match underdogs respond to financial incentives. Results on *GPM* and *WPM* retain expected directionality and are statistically significant at the 1% level, affirming that increasing guaranteed payout and/or decreasing winning payout increases the odds that underdogs will quit. Results on *RankDiff, BetterRank,* and *BetterRank*RankDiff* remain insignificant with coefficients close to 1, likely for reasons discussed earlier.

Coefficients on the interaction terms *BetterRank*GPM, BetterRank*GPo, BetterRank*WPo,* and *BetterRank*WPM* have expected directionality and are significant at the 5% level, except for *BetterRank*WPM*, which is significant at the 10% level. The significance of these results suggests that the distinction of being the match underdog or favorite *does* influence the role of incentives on odds of quitting. Interestingly, match underdogs may be more influenced than favorites by the role incentives in quitting behaviors; they may exhibit more susceptibility to quitting than favorites when guaranteed prizes increase and less susceptibility to quitting when winning prizes increase.

7.1.3 Model III: Base Model with Unseeded Player Interactions

Interacting the *Seeded* variable with the prize incentive variables allows this thesis to account for the possibility that being seeded or unseeded in a tournament changes the way prize

[16] A chi-squared test in the data section suggests that match underdogs quit more than match favorites.

incentives influence quitting behaviors. It differs from the *BetterRank* interaction regression because it gauges the more general, longer term phenomenon of how incentives influence player quitting behaviors when players compare themselves not to their immediate opponent, but to the overall field.[17] Recall that the interaction effect is included in a separate model because there is collinearity between being a match favorite and being seeded. The Pseudo R^2 value of 0.057 and the chi-squared statistic value of 376.18 (with 187 degrees of freedom and a p-value of 0.000) suggest that Model III is statistically significant and has similar explanatory power and goodness-of-fit as both Models I and II.

Results of Model III again reinforce that financial incentives matter for player quitting behaviors. Coefficients on *GPM* and *WPM* retain expected directionality and are significant at the 5% and 10% levels, respectively. Similar to the discussion in Model II, coefficients on *GPM* and *WPM* represent the influence of financial incentives on quitting behavior among unseeded players only. Results indicate that increasing the guaranteed prize increases the unseeded player's odds of quitting and that increasing the winning prize decreases the unseeded player's odds of quitting. Coefficients of the skill heterogeneity variables remain statistically insignificant and are close to 1, likely for reasons discussed earlier.

Coefficients on the interaction terms themselves are statistically insignificant, except for *Seeded*GPM*, which is statistically significant at the 10% level with expected directionality. Thus, results are generally inconclusive. This seems to make sense. The seeded vs. unseeded distinction is more generalized than the match favorite vs. underdog distinction; it is entirely possible that an unseeded player is still the favorite of a given match.

---

[17] Similarly, a chi-squared test in the data section suggests that unseeded players quit more than seeded players.

7.2 Grand Slam Dataset

Running the models on the isolated Grand Slam dataset is important for two reasons. First, correlational analyses in the Data section show that players exhibit higher rates of quitting in the most lucrative tournaments on tour. Running the models on the isolated Grand Slam dataset allows this thesis to analyze how incentives influence quitting behaviors when said incentives are most exaggerated. In addition, the confounding effect between ranking point incentives and financial incentives discussed earlier is avoided because Grand Slams offer consistent guaranteed and winning ranking point distributions. So, running the model on the Grand Slam dataset allows this thesis to better identify the incentive effect of guaranteed and winning ranking points on quitting behaviors.

Before interpreting results, it is important to note a couple of general factors. First, restricting the data to only Grand Slam tournaments limits the number of observations to 6,185. Because retirements are rare occurrences, this substantially limits the number of quitting instances to analyze, adding noise that could distract from results. Accordingly, there are generally few statistically significant results in the Grand Slam-only cohort. Second, the underlying economic logic for how incentives impact quitting behaviors in Grand Slams is no different from the data in general. The only expected differences are which types of players are more likely to quit and how the consistent ranking point distribution impacts quitting behaviors.

7.2.1 Model I: Base Model

The first column of the Grand Slam dataset includes output from Model I without fixed effects estimators.[18] Similar to the earlier discussion, including fixed effects estimators in the

[18] Just as in the combined dataset, including player fixed effects in the regression causes Stata to drop all players who have never quit, so the model without fixed effects estimators has more observations than the corresponding model with fixed effects estimators.

model increases the Pseudo R^2 from 0.013 to 0.068. However, while the chi-squared statistic of 26.94 in the model without fixed effects is statistically significant (with 7 degrees of freedom and a p-value of 0.000), the chi-squared statistic of 129.15 in the model with fixed effects estimators is not (with 133 degrees of freedom and a p-value of 0.58). The lack of significance is likely due to increased degrees of freedom added to the model when accounting for the fixed effects. As before, for economic reasons this section only discusses Model I results with fixed effects estimators.

Results of the base model are not statistically significant, likely due to the sample size problems discussed earlier. So, outcomes must be interpreted with caution. That being said, results suggest that incentives continue to influence quitting behaviors, as evidenced by expected directionality on coefficients of all prize incentive variables. Of particular note is that, in accordance with the earlier discussion, coefficients on ranking point incentive variables *GPo* and *WPo* now exhibit expected directionality, suggesting that ranking point incentives, as with direct financial incentives, may influence quitting behaviors. Increasing guaranteed ranking points could increase odds of quitting; increasing winning ranking points could decrease odds of quitting. The discussion regarding skill heterogeneity variables is also reinforced. *RankDiff* and *BetterRank*RankDiff* remain statistically insignificant with coefficients close to 1.

That being said, while *BetterRank* remains statistically insignificant, its coefficient exhibits expected directionality, and its magnitude is notably lower than its corresponding value in the combined dataset. This acts as additional validation for the hypothesis that, while the exact relationship between skill heterogeneity and quitting behaviors is unclear, the simple distinction of "being the favorite" may influence quitting behaviors.

7.2.2 Model II: Base Model with Match Underdog Interaction

Model II features a Pseudo R^2 value of 0.076 in the Grand Slam dataset. Just as in Model I, its chi-squared statistic of 143.49 is statistically insignificant (with 137 degrees of freedom and a p-value of 0.335), likely for similar reasons.

Results of Model II on the isolated grand slam dataset largely reinforce results already discussed in the combined data section, though experience a similar lack of statistical significance as Model I. That being said, there are three notable exceptions. First, in this iteration of the model results for *BetterRank* are statistically significant at the 5% level, with continued as-expected directionality. This is the only model in which results on *BetterRank* are statistically significant, offering support for the hypothesis that being the match favorite reduces the odds of quitting during Grand Slam tournaments. The second and third exceptions build upon the observed trend on *BetterRank*: coefficients of the interaction of *BetterRank* with the ranking incentive variables *GPo* and *WPo* are statistically significant at the 1% level with expected directionality. These results offer further support that match underdogs and match favorites may respond differently to equivalent changes in ranking point incentives within Grand Slam tournaments. Match underdogs may be more influenced than favorites to quit when guaranteed and winning ranking point prizes change.

7.2.3 Model III: Base Model with Unseeded Player Interaction

Model III features a Pseudo-R^2 value of 0.074 and a chi-squared statistic of 139.65 (with 187 degrees of freedom and a p-value 0.445), indicating similar significance and predictive power as Models I and II.

Results of Model III are largely statistically insignificant, save *GPM* which is statistically significant at the 10% level and exhibits expected directionality. This seems to reinforce the earlier discussion that the match underdog vs. favorite distinction is potentially a more useful measure of how players of different skill respond to incentives when considering quitting, given the general nature of the seeded vs. unseeded classification. That being said, results in general exhibit expected directionality and are consistent with previous results on incentive, ranking heterogeneity, and interaction variables, offering further support for the earlier discussion of Model III and of general results.

8. Conclusion and Further Research

This thesis studies the effect of incentive structures on quitting behaviors in men's professional tennis tournaments. Publicly available data published by the Association of Tennis Professionals (ATP) makes it possible to test whether specific incentive factors encourage players to quit from tournaments, what kinds of players are most likely to quit from tournaments, and how specific subgroups of players respond to quitting incentives differently.

Foremost, consistent with Rosen's hypotheses, results suggest that incentives influence quitting behaviors. Players are more likely to quit from tennis tournaments when their guaranteed loser's prize increases and are less likely to quit when the winning prize increases. Results of the isolated Grand Slam cohort suggest that players may respond similarly to indirect financial incentives, as measured by ranking points. Generally insignificant results on measurements of skill heterogeneity indicate that the relationship between difference in skill and quitting likelihood is less clear, though being the match underdog may increase odds of quitting.

This paper makes the additional conclusion that certain types of players – the worse-ranked players in individual matches and, possibly, "unseeded" players in tournaments – may be particularly susceptible to the influence of prize incentives on quitting behaviors. These results suggest that emphasizing winning pay over guaranteed pay could increase tournament efficiency and player effort exertion by not only enhancing optimal performance (as discussed at length in precedent literature), but by reducing harmful quitting behaviors, particularly among less-skilled players.

One can apply the findings of this paper to the labor market by viewing a professional tennis tournament as a proxy for how workers interact in promotion and bonus competitions. In particular, if one views the guaranteed prize as the starting salary, the winning prize as the promotion salary or bonus, and relative player skill as various measurements of worker performance, it is easy to see how the incentive structures that influence whether a player will quit a tennis match can mirror those that influence whether an employee will forego bonus promotion or bonus considerations. If one is inclined to agree with the analogous characterizations presented in the labor market, then offering sufficiently high starting salaries to employees may discourage employees from exerting the effort necessary to be considered for promotions or bonuses (in not going above and beyond to demonstrate their readiness for a higher-level position, they are effectively "quitting" bonus or promotion considerations). This would likely concern leaders interested in the future success of the company, particularly if the company hires entry-level employees with the intention of training them for leadership roles.[19] The logical application is that placing a greater percent of an employee's pay on relative performance (the "winning prize" component) could reduce inefficient work behavior and increase productivity of the firm as a whole.

The preliminary findings that "underdog" populations could be more susceptible to financial incentives in tennis tournaments may also be applied to the labor market. Many firms have "star performers" who other employees could view as more likely to receive a bonus or promotion than them. Results of this paper may be extended to suggest that employees who view themselves as underdogs versus the star performer (or relative to other employees in their division) could be most susceptible to engaging in inefficient work behaviors. But results also suggest that underdog employees could be especially receptive to exerting effort for bonus consideration if a greater percent of their pay is tied to relative performance-based compensation as opposed to a guaranteed salary.

[19] This is the case in many professional service firms with a "Partner track."

Of course, results of this paper are preliminary and not definitively conclusive. A number of opportunities exist to expand empirical understanding of tournament theory's explanation of quitting behaviors, which could either reinforce or challenge this work's conclusions. First, extending the model beyond 2017 would allow an economist to evaluate whether institutional rule changes implemented by the ATP (discussed in the background section) can reduce the impact of the incentive effect on quitting behaviors. Because the rule-change is so recent, exploring it is outside the scope of this paper. Second, recall that the four-member Grand Slam board changed the incentive structure of its tournaments in 2018 to allow players to pull out before tournaments and keep 50% of their first round prize money. It would be interesting to run the regressions created in this thesis on Grand Slam tournaments in 2018 and beyond, because it would allow for a comparison of two different types of adjustments for compensation for pulling out before tournaments: one in which players keep all of the first round loser's prize on two occasions (non-Grand Slams), and another in which players keep half of the first round loser's prize for all occasions (Grand Slams). The results of these analyses may shed light on how the tournament director could best reorganize tournament incentives to reduce the influence of finances on quitting behaviors.

Likewise, additional work could incorporate different measurements of incentives on player quitting behaviors. While most precedent literature analyzed marginal payoff for advancing one round in a tournament, different measures of incentives have been used to study the effect of finances on quitting behaviors. Ivankovic (2007) and Silverman & Seidel (2011) studied how the discounted present value of potential earnings for advancing through all subsequent tournament rounds impacts effort exertion. Future work could incorporate this type of measurement in a study of quitting behaviors.

Finally, similar to a point made by Silverman & Seidel (2011), future work could be undertaken to link quitting behavior in the professional sports arena with quitting behaviors in the corporate world. It is always tenuous to extrapolate the findings of a single, isolated system (the ATP tour) to the very generalized corporate world. Because this paper is among the first to apply tournament theory to quitting behaviors (and because perceived analogous behaviors in the labor market could have significant impacts on employee livelihood), results found here should be extrapolated to the labor market with caution. Similar studies analyzed in separate sports or related activities may help fill out the picture that this thesis sketches, providing labor economists with a more comprehensive view of how incentive structures and related factors could influence quitting behaviors.

Works Cited

ATP World Tour (n.d.). Retrieved from https://www.atpworldtour.com

ATP Official Rulebook (2017). *ATP*. Retrieved from https://www.atpworldtour.com/-/media/files/rulebook/2017/2017-atp-rulebook_3feb17.pdf

Connelly, B., Tihanyi, L., Crook, T.R., Gangloff, K.A. (2014). Tournament Theory: Thirty Years of Contests and Competitions. *Journal of Management, 40* (1), 16-47.

Corral, J.D., Pireto-Rodriguez, J. (2010). Are differences in ranks good predictors for Grand Slam tennis matches? *International Journal of Forecasting*, *26* (3), 551-563.

DeVaro, J. (2006). Strategic promotion tournaments and worker performance. Retrieved from Cornell University, School of Industrial and Labor Relations site: http://digitalcommons.ilr.cornell.edu/articles/108/

Ehrenberg, R. G. & Bognanno, M. L. (1990). Do tournaments have incentive effects? *Journal of Political Economy, 98* (6), 1307-1324.

Ekinci, E. and Waldman, M. (2015). Bonuses and Promotion Tournaments: Theory and Evidence. Retrieved from Universidad Carlos III de Madrid Department of Business Administration and Cornell University Johnson Graduate School of Management: http://www.sole-jole.org/16016.pdf

Geyer, H. (2010). Quit Behavior of Professional Tennis Players. *Journal of Sports Economics*, *11* (1) 89-99.

Gilsdorf, K., & Sukhatme, V. (2008). Testing Rosen's Sequential Elimination Tournament Model: Incentives and Player Performance in Professional Tennis. *Journal of Sports Economics*, *9* (3), 287-303.

Grund, C., Höcker, J, and Zimmermann, S. (2010). Risk Taking Behavior in Tournaments: Evidence from the NBA. *IZA Discussion Paper* (*4812)*. Retrieved from https://ssrn.com/abstract=1570430

Hvide, H. (2002). Tournament Rewards and Risk Taking. *Journal of Labor Economics, 20* (4), 877-898.

Henderson, A. D., Fredrickson, J. W. (2001). TMT coordination needs and the CEO pay gap: A competitive test of economic and behavioral views. The *Academy of Management Journal*, *44* (1), 96-117.

Ivankovic, M. (2007). The tournament model: an empirical investigation of the ATP Tour. *Zb. rad. Ekon. fak. Rij*, *25* (1), 83-111.

Knoeber, C. R., Thurman, W. N. (1994). Testing the theory of tournaments: An empirical analysis of broiler production. *Journal of Labor Economics*, *12* (1), 155-179.

Lazear, P.E., Rosen, S. (1981). Rank-Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy*, 89 (5), 841-864.

Lazear, E. P. (1989). Pay Equality and Industrial Politics. *Journal of Political Economy 97*, (3) 561-580.

Lallemand, T., Plasman, R., Rycx, F. (2008). Women and competition in elimination tournaments. *Journal of Sports Economics*, *9* (1), 3-19.

Lee, J. (2004). Prize and Risk-Taking Strategy in Tournaments: Evidence from Professional Poker Players. *IZA Discussion Paper (1345)*. Retrieved from SSRN: https://ssrn.com/abstract=603525

Nalebuff, B.J., Stiglitz, J.E. (1983). Prizes and Incentives: Toward a General Theory of Compensation and Competition. *Bell Journal of Economics 14* (1), 21-43.

OFX (n.d.). Historical Exchange Rates. Retrieved from https://www.ofx.com/en-us/forex-news/historical-exchange-rates/

O'Roark, B. J., Wood, C. W., Demblowski, B. (2012). Tournament Chasing NASCAR Style: Driver Incentives in Stock Car Racing's Playoff Season. *Eastern Economic Journal*, *38* (1), 1-17

Ovaska, T., Sumell, A. (2014). Who Has the Advantage? An Economic Exploration of Winning in Men's Professional Tennis. *The American Economist*, 59 (1), 34-51.

Pro Tennis Live (n.d.). Draws and Prize Payouts for Men's Professional Tennis Tournaments. Retrieved from https://www.protennislive.com/

Rosen, S. (1986). Prizes and Incentives in Elimination Tournaments. *American Economic Review*, *76* (4), 701-715.

Rothenberg, B (2018, January 16). Retirements Are Down at the Australian Open. Is Money the Reason? *The New York Times*. Retrieved from www.nytimes.com/2018/01/16/sports/tennis/australian-open-retirements.html.

Silverman, J., Seidel, S. (2011). "Incentives in Professional Tennis: Tournament Theory and Intangible Factors." *Duke University Honors Thesis*. Retrieved from https://sites.duke.edu/econhonors/2011/04/03/incentives-in-professional-tennis-tournament-theory-and-intangible-factors/

Sunde, U. (2003). "Potential, Prizes, and Performance: Testing Tournament Theory with Professional Tennis Data." *IZA Discussion Paper Series (947)*, 1-28.

Sunde, U. (2009). Heterogeneity and performance in tournaments: a test for incentive effects using professional tennis data. *Applied Economics*, *41* (29), 3199-3208.

US Bureau of Labor Statistics (n.d.). CPI Inflation Calculator. Retrieved from https://www.bls.gov/data/inflation_calculator.htm

Waldstein, D. (2017, November 21). Grand Slam Events Take Steps to Curb Mid-Match Retirements. The New York Times, *The New York Times*. Retrieved from www.nytimes.com/2017/11/21/sports/tennis/grand-slam-board-rules-changes.html.

Wikipedia (n.d.). Round-By-Round Payout Information for Grand Slam Tournaments. Retrieved from https://www.wikipedia.org/

Wimbledon Championships Official Prize Money (2017). *The Championships, Wimbledon.* Retrieved from http://www.wimbledon.com/pdf/prize_money_2017.pdf

Appendix

Table 1: Variables and Definitions

VARIABLE	DEFINITION
RET	Dependent variable for whether the player quits (1 if quits, 0 if not)
GPM	The prize money guaranteed to a player for making it to a given round in a tournament (USD, thousands)
WPM	The guaranteed prize money plus the marginal prize money gain awarded to a player should they win a given match (USD, thousands)
GPO	The ranking points guaranteed to a player for making it to a given round in a tournament
WPO	The guaranteed ranking points plus the marginal point gain awarded to a player should they win a given match
RANKDIFF	The difference in year-end ranking between match competitors in the year prior to the year of the given match
BETTERRANK	Binomial dummy variable: 1 = better year-end ranking (favorite), 0 = worse year-end ranking (underdog)
RANKDIFFBETTER	Interaction term multiplying ranking difference by a binomial value of whether the player is ranked better or worse (1 if better, 0 if worse)
SEEDED	Binomial dummy variable: 1 = player is seeded in the tournament, 0 = player is unseeded in the tournament
PCODE	Player fixed effects estimator; encoded string of player name
SCODE	Court surface fixed effects estimator; encoded string of court surface (1 = clay, 2 = grass, 3 = hard)
DRAWSIZE	Tournament fixed effects estimator; can be 32, 64, or 128
ROUND	Fixed effects estimator for match stage; ranges from 1-7
YEAR	Fixed effects estimator for the year a match within a tournament takes place; ranges from 2007-2017

Table 2: Variable & Interaction Effect Predictions

VARIABLE	EXPECTED DIRECTION	HYPOTHESIS
RET	N/A	Dependent variable
GPM	>1	Increasing guaranteed prize money increases incentive to quit
WPM	<1	Increasing winning prize money decreases incentive to quit
GPO	>1	Increasing guaranteed ranking points increases incentive to quit
WPO	<1	Increasing winning ranking points decreases incentive to quit
RANKDIFF	>1	Higher skill heterogeneity increases incentive for a player (usually underdog) to quit
BETTERRANK	<1	Better ranked payers have higher chances of winning, so are less likely to quit
RANKDIFF*BETTERRANK	<1	Underdogs are more likely to quit in matches with high skill difference
BETTERRANK*GPM	<1	Underdogs are more influenced by financial incentives than favorites
BETTERRANK*WPM	>1	Underdogs are more influenced by financial incentives than favorites
BETTERRANK*GPO	<1	Underdogs are more influenced by ranking incentives than favorites
BETTERRANK*WPO	>1	Underdogs are more influenced by ranking incentives than favorites
SEEDED	<1	Seeded players have higher probability of winning, so are less likely to quit
SEEDED*GPM	<1	Unseeded players are more influenced by financial incentives than favorites
SEEDED*WPM	>1	Unseeded players are more influenced by financial incentives than favorites
SEEDED*GPO	<1	Unseeded players are more influenced by ranking incentives than favorites
SEEDED*WPO	>1	Unseeded players are more influenced by ranking incentives than favorites
PCODE	N/A	Fixed Effects Estimator
SCODE	N/A	Fixed Effects Estimator
DRAWSIZE	N/A	Fixed Effects Estimator
ROUND	N/A	Fixed Effects Estimator
YEAR	N/A	Fixed Effects Estimator

Table 3: Descriptive Statistics for Combined Dataset

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Retire	29,547	.023	.151	0	1
WPM (thousands)	29,547	72.05	124.85	5.94	1825
GPM (thousands)	29,547	38.4028	61.65	2.20	920
GPo	29,547	51.49	86.38	0	720
WPo	29,547	106.89	146.47	10	1200
RankDiff	29,547	57.03	62.98	1	492
BetterRank	29,547	.561	0.500	0	1
RankDiff*BetterRank	29,547	35.57	60.40	0	492
BetterRank*GPM	29,547	20.67	47.99	0	920
BetterRank*WPM	29,547	38.60	95.97	0	1825
BetterRank*GPo	29,547	27.16	66.62	0	720
BetterRank*WPo	29,547	57.03	116.81	0	1200
Seeded	29,547	0.569	0.495	0	1
Seeded*GPM	29,547	27.89	62.50	0	920
Seeded*WPM	29,547	53.16	126.12	0	1825
Seeded*GPo	29,547	39.73	86.48	0	720
Seeded*WPo	29,547	78.33	151.12	0	1200

Table 4: Descriptive Statistics for Grand Slam Dataset

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Retire	6,185	.036	.185	0	1
WPM	6,185	106.68	115.00	23.89	920
GPM	6,185	59.73	58.96	14.95	470
GPo	6,185	56.07	74.39	5	360
WPo	6,185	123.05	140.42	35	720
RankDiff	6,185	65.41	67.28	1	476
BetterRank	6,185	.570	.50	0	1
RankDiff*BetterRank	6,185	41.77	64.68	0	476
BetterRank*GPM	6,185	31.76	49.04	0	470
BetterRank*WPM	6,185	56.17	92.45	0	920
BetterRank*GPo	6,185	28.30	56.35	0	360
BetterRank*WPo	6,185	123.05	140.42	0	720
Seeded	6,185	.583	.493	0	1
Seeded*GPM	6,185	42.014	64.21	0	470
Seeded*WPM	6,185	76.24	123.03	0	920
Seeded*GPo	6,185	56.07	74.34	5	360
Seeded*WPo	6,185	90.81	148.87	0	720

Table 5: Regression Results, Combined Dataset

VARIABLES	(1) Model I, No Fixed Effects	(2) Model I, Fixed Effects	(3) Model II: Match Underdog Interaction	(4) Model III: Unseeded Player Interaction
GPM	1.021**	1.018*	1.032***	1.051**
(USD, thousands)	(0.00869)	(0.0103)	(0.0124)	(0.0219)
WPM	0.989**	0.989*	0.983***	0.977*
(USD, thousands)	(0.00470)	(0.0055)	(0.00654)	(0.0119)
GPo	0.991**	0.993	1.001	0.986
	(0.00423)	(0.00636)	(0.00749)	(0.00924)
WPo	1.005**	1.006*	1.001	1.008
	(0.00256)	(0.00339)	(0.00408)	(0.00564)
BetterRank	0.799**	0.893	0.797	0.923
	(0.0853)	(0.0998)	(0.131)	(0.105)
RankDiff	0.998*	0.998	0.998	0.999
	(0.00113)	(0.00117)	(0.00118)	(0.00119)
BetterRank*RankDiff	1.001	1.000	1.000	1.000
	(0.00143)	(0.00145)	(0.00147)	(0.00147)
BetterRank*GPM	[N/A]	[N/A]	0.965**	[N/A]
	[N/A]	[N/A]	(0.0172)	[N/A]
BetterRank*WPM	[N/A]	[N/A]	1.018*	[N/A]
	[N/A]	[N/A]	(0.0101)	[N/A]
BetterRank*GPo	[N/A]	[N/A]	0.980**	[N/A]
	[N/A]	[N/A]	(0.00874)	[N/A]
BetterRank*WPo	[N/A]	[N/A]	1.012**	[N/A]
	[N/A]	[N/A]	(0.00546)	[N/A]
Seeded	[N/A]	[N/A]	[N/A]	1.047
	[N/A]	[N/A]	[N/A]	(0.177)
Seeded*GPM	[N/A]	[N/A]	[N/A]	0.957*
	[N/A]	[N/A]	[N/A]	(0.0219)
Seeded*WPM	[N/A]	[N/A]	[N/A]	1.018
	[N/A]	[N/A]	[N/A]	(0.0136)
Seeded*GPo	[N/A]	[N/A]	[N/A]	1.008
	[N/A]	[N/A]	[N/A]	(0.00999)
Seeded*WPo	[N/A]	[N/A]	[N/A]	0.998
	[N/A]	[N/A]	[N/A]	(0.00624)
Constant	0.0247***	0.00874***	0.00882***	0.00782***
	(0.00238)	(0.00551)	(0.00559)	(0.00500)
Player Fixed Effects[20]	No	Yes	Yes	Yes
Tourn/Match FE[21]	No	Yes	Yes	Yes
Observations	32,097	29,547	29,547	29,547
Pseudo R^2	0.005	0.055	0.056	0.057
Chi^2 Statistic (df)	31.41 (7)	360.36 (182)	368.87 (186)	376.18 (187)
P-Value	0.000	0.000	0.000	0.000
Log Likelihood	-3316.31	-3094.06	-3089.81	-3086.15

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[20] Includes player name
[21] Includes year, draw size, surface, and round

Table 6: Regression Results, Grand Slam Dataset

VARIABLES	(1) Model I, No Fixed Effects	(2) Model I, Fixed Effects	(3) Model II: Match Underdog Interaction	(4) Model III: Unseeded Player Interaction
GPM	1.020*	1.013	1.026	1.059*
(USD, thousands)	(0.0114)	(0.0147)	(0.0181)	(0.0341)
WPM	0.990	0.992	0.985	0.970
(USD, thousands)	(0.00633)	(0.00802)	(0.0101)	(0.0184)
GPo	0.982*	1.077	1.099	1.062
	(0.0101)	(0.0892)	(0.0941)	(0.0910)
WPo	1.009*	0.983	0.974	0.993
	(0.00561)	(0.0240)	(0.0243)	(0.0276)
BetterRank	0.598***	0.783	0.365**	0.889
	(0.117)	(0.163)	(0.160)	(0.191)
RankDiff	0.999	0.999	0.999	0.999
	(0.00168)	(0.00175)	(0.00177)	(0.00180)
BetterRank*RankDiff	1.001	1.001	1.002	1.001
	(0.00221)	(0.00226)	(0.00230)	(0.00227)
BetterRank*GPM	[N/A]	[N/A]	0.969	[N/A]
	[N/A]	[N/A]	(0.0277)	[N/A]
BetterRank*WPM	[N/A]	[N/A]	1.018	[N/A]
	[N/A]	[N/A]	(0.0165)	[N/A]
BetterRank*GPo	[N/A]	[N/A]	0.918***	[N/A]
	[N/A]	[N/A]	(0.0269)	[N/A]
BetterRank*WPo	[N/A]	[N/A]	1.046***	[N/A]
	[N/A]	[N/A]	(0.0167)	[N/A]
Seeded	[N/A]	[N/A]	[N/A]	1.197
	[N/A]	[N/A]	[N/A]	(0.561)
Seeded*GPM	[N/A]	[N/A]	[N/A]	0.946
	[N/A]	[N/A]	[N/A]	(0.0340)
Seeded*WPM	[N/A]	[N/A]	[N/A]	1.028
	[N/A]	[N/A]	[N/A]	(0.0217)
Seeded*GPo	[N/A]	[N/A]	[N/A]	1.026
	[N/A]	[N/A]	[N/A]	(0.0311)
Seeded*WPo	[N/A]	[N/A]	[N/A]	0.986
	[N/A]	[N/A]	[N/A]	(0.0173)
Constant	0.0274***	0.0168***	0.0161***	0.00985***
	(0.00532)	(0.0204)	(0.0197)	(0.0125)
Player Fixed Effects[22]	No	Yes	Yes	Yes
Tourn/Match FE[23]	No	Yes	Yes	Yes
Observations	9,222	6,185	6,185	6,185
Pseudo R^2	0.013	0.068	0.076	0.074
Chi^2 Statistic (df)	26.94 (7)	129.15 (133)	143.49 (137)	139.65 (138)
P-Value	0.000	0.578	0.335	0.445
Log Likelihood	-1025.74	-885.44	-878.27	-880.19

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[22] Includes player name
[23] Includes year, draw size, surface, and round

Table 7: Regression Results, Combined Dataset (Using RankDiff^2 for Heterogeneity)

VARIABLES	(1) Model I: Base Model	(2) Model II: Match Underdog Interaction	(3) Model III: Unseeded Player Interaction
GPM	1.018*	1.032***	1.050**
(USD, thousands)	(0.0103)	(0.0124)	(0.0219)
WPM	0.990*	0.983***	0.977*
(USD, thousands)	(0.00552)	(0.00655)	(0.0119)
GPo	0.993	1.001	0.986
	(0.00636)	(0.00750)	(0.00926)
WPo	1.006*	1.001	1.008
	(0.00339)	(0.00409)	(0.00565)
BetterRank	0.897	0.805	0.931
	(0.0819)	(0.119)	(0.0865)
RankDiff^2	**1.000**	**1.000**	**1.000**
	(4.19e-06)	**(4.22e-06)**	**(4.18e-06)**
BetterRank*RankDiff^2	**1.000**	**1.000**	**1.000**
	(5.05e-06)	**(5.07e-06)**	**(5.03e-06)**
BetterRank*GPM	[N/A]	0.965**	[N/A]
	[N/A]	(0.0172)	[N/A]
BetterRank*WPM	[N/A]	1.018*	[N/A]
	[N/A]	(0.0101)	[N/A]
BetterRank*GPo	[N/A]	0.980**	[N/A]
	[N/A]	(0.00873)	[N/A]
BetterRank*WPo	[N/A]	1.012**	[N/A]
	[N/A]	(0.00546)	[N/A]
Seeded	[N/A]	[N/A]	1.022
	[N/A]	[N/A]	(0.172)
Seeded*GPM	[N/A]	[N/A]	0.958*
	[N/A]	[N/A]	(0.0220)
Seeded*WPM	[N/A]	[N/A]	1.018
	[N/A]	[N/A]	(0.0136)
Seeded*GPo	[N/A]	[N/A]	1.008
	[N/A]	[N/A]	(0.01000)
Seeded*WPo	[N/A]	[N/A]	0.998
	[N/A]	[N/A]	(0.00625)
Constant	0.00840***	0.00854***	0.00741***
	(0.00528)	(0.00538)	(0.00472)
Player Fixed Effects[24]	Yes	Yes	Yes
Tourn/Match FE[25]	Yes	Yes	Yes
Observations	29,547	29,547	29,547
Pseudo R^2	.055	.056	.057
Chi^2 Statistic (df)	358.33 (182)	366.81 (186)	374.41 (187)
P-Value	0.000	0.000	0.000
Log Likelihood	-3095.08	-3090.84	-3087.04

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[24] Includes player name
[25] Includes year, draw size, surface, and round

Table 8: Regression Results, Grand Slam Dataset (Using RankDiff^2 for Heterogeneity)

VARIABLES	(1) Model I: Base Model	(2) Model II: Match Underdog Interaction	(3) Model III: Unseeded Player Interaction
GPM	1.014	1.026	1.060*
(USD, thousands)	(0.0147)	(0.0181)	(0.0340)
WPM	0.992	0.985	0.969*
(USD, thousands)	(0.00801)	(0.0101)	(0.0184)
GPo	1.078	1.099	1.062
	(0.0893)	(0.0942)	(0.0911)
WPo	0.983	0.974	0.993
	(0.0241)	(0.0244)	(0.0276)
BetterRank	0.815	0.384**	0.910
	(0.138)	(0.159)	(0.159)
RankDiff^2	**1.000**	**1.000**	**1.000**
	(5.15e-06)	**(5.16e-06)**	**(5.05e-06)**
BetterRank*RankDiff^2	**1.000**	**1.000**	**1.000**
	(6.45e-06)	**(6.51e-06)**	**(6.38e-06)**
BetterRank*GPM	[N/A]	0.969	[N/A]
	[N/A]	(0.0276)	[N/A]
BetterRank*WPM	[N/A]	1.018	[N/A]
	[N/A]	(0.0165)	[N/A]
BetterRank*GPo	[N/A]	0.918***	[N/A]
	[N/A]	(0.0269)	[N/A]
BetterRank*WPo	[N/A]	1.046***	[N/A]
	[N/A]	(0.0167)	[N/A]
Seeded	[N/A]	[N/A]	1.172
	[N/A]	[N/A]	(0.547)
Seeded*GPM	[N/A]	[N/A]	0.946
	[N/A]	[N/A]	(0.0339)
Seeded*WPM	[N/A]	[N/A]	1.029
	[N/A]	[N/A]	(0.0217)
Seeded*GPo	[N/A]	[N/A]	1.027
	[N/A]	[N/A]	(0.0310)
Seeded*WPo	[N/A]	[N/A]	0.986
	[N/A]	[N/A]	(0.0173)
Constant	0.0155***	0.0147***	0.00944***
	(0.0188)	(0.0180)	(0.0119)
Player Fixed Effects[26]	Yes	Yes	Yes
Tourn/Match FE[27]	Yes	Yes	Yes
Observations	6,185	6,185	6,185
Pseudo R^2	0.068	0.076	0.073
Chi^2 Statistic (df)	129.09 (133)	143.37 (137)	139.95 (138)
P-Value	0.580	0.338	0.438
Log Likelihood	-885.47	-878.33	-880.04

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[26] Includes player name
[27] Includes year, draw size, surface, and round

Table 9: Correlation Matrix Among Independent Variables

	WPM	GPM	GPo	WPo	RankDiff	BetterRank	Seeded
WPM	1.00						
GPM	1.00	1.00					
GPo	0.92	0.91	1.00				
WPo	0.92	0.91	1.00	1.00			
RankDiff	-0.13	-0.13	-0.17	-0.17	1.00		
BetterRank	0.00	0.00	0.00	0.00	0.00	1.00	
Seeded	0.16	0.16	0.21	0.21	0.14	0.22	1.00

Addendum 1: Data Cleaning Process

The initial task involved the construction of the dataset, which included determining which tournaments to analyze and the time horizon from which to choose tournaments. The latter informed the former. Ultimately, the dataset reflects an 11-year timeline: every data point is drawn from a match that occurred in the years 2007-2017. This time frame was settled upon because it captures what, in the opinion of this author, is a sufficiently large and robust sample of tennis matches while being manageable to collect. Due to the manual nature of the data collection process, obtaining more data points would entail a significant time commitment (as reference, it took roughly 50 hours to collect and clean the dataset in its current state). It was then necessary to decide upon tournaments from which to analyze individual matches. It was decided that the dataset would include all yearly tournaments that successfully ran throughout the duration of the 11-year time horizon (i.e., all of the tournaments were not discontinued during this time frame). The sample covers all of the three major court surfaces in tennis (hard, grass, and clay)[28] and all tournament classifications (ATP 250, ATP 500, Masters 1,000, and Grand

[28] Tournaments that took place on the carpet surface are not included in this analysis.

Slam).[29] They also represent a robust sampling of prize money; the dataset includes tournaments with as little as $380,000 in prize money and with more than $20M in prize money.

The construction and data scraping process was relatively robust. As the ATP tour reports match-by-match scores and results in table formats on their website, the "importhtml" function on Google Sheets was used to pull relevant information regarding the identified tournaments from the ATP website into a spreadsheet. The excel file was then reformatted largely manually by adjusting the imported information so that, left to right, its columns were organized according to relevant independent and dependent variables. Table 1 includes definitions of all independent variables included in this thesis's model.

As prize money distributions of individual tournaments are not recorded in the ATP's publicly available dataset, prize money information on a tournament-by-tournament basis was appended to the dataset. They were pulled separately from two different sources. Information for prize money information, round-by-round, was pulled from each of the 4 Grand Slam Tournament's Wikipedia pages (n.d.) because the tournaments archive their payout information each year, making past results difficult to obtain from their websites.[30] To access the information, enter [desired year] and [Grand Slam tournament name] into the Wikipedia search bar. All relevant prize distributions are included in the "Prize Money" section. Information for all non-Grand Slam tournaments was pulled from the archives of their official draws, which are stored in www.protennislive.com (n.d.). The actual information was transferred manually from its origin data source to the created dataset.

---

[29] Prior to 2009, professional tournaments were not classified in the manner of ATP 250 vs. ATP 500, and so on. This is reflected in the point and prize money distributions of tournaments in 2007 and 2008.
[30] I verified these values by cross-referencing them with Figures I found in press releases from the tournaments and news articles from third-party analysts.

Ranking information (including year-over-year rank for a given player and the number of tournaments he competed in per year) are also available publicly on the ATP's website, but are stored separately from individual match information. Thus, year-over-year rank and number of tournaments played for each player were appended in a similar manner as prize money data was. The initial process of construction described here resulted in the 19,089 data points shown in the first row of Figure 4. Due to standard formatting of the ATP data, these nearly 20,000 data points were organized by match, with information about the match's winner and loser in different columns within the same row.

Because this thesis is interested in exploring the behaviors of *players* who engage in retirements (as opposed to *matches in which* retirements occur), it was necessary to reorganize the dataset so that each row represented a single player's actions in a match within a tournament.[31] Accordingly, the "Winner" and "Loser" format of the raw data was transformed into a single column, "Player." In doing so, several player-specific factors were unlocked for analysis, including: player rank, Qualifier status, Wild Card status, Unseeded Status, the number of prior tournaments played, whether or not the given player is the better ranked player in the match, and so on. This doubled the size of the dataset, and corresponds to the second row of Figure 4 showing 38,178 observations.

Unfortunately, prize money distribution for some tournaments was unavailable on the internet, either due to inability to access archived draws or another reason unbeknownst to this author. Because prize money is a key factor in a player's decision of whether to retire, all tournaments without publicly available prize money distributions were dropped from the dataset.

[31] In the prior format, each observation was recorded on a match level, which by definition includes two players.

This corresponds to the third row of Figure 4 showing 37,614 observations (representing 18,807 matches).

Not all matches competed on the ATP tour are between two ranked opponents. For any number of reasons, a player in a given match may be unranked. Because a player's rank is an important component of the theoretical framework behind this analysis, all observations from unranked players were removed from the dataset, resulting in 36,536 remaining observations.

In the sample dataset, no retirements were observed in the final round of the largest international tournaments (Round 7). Because of perfect prediction of behaviors in Round 7, Stata dropped all Round 7 matches when calculating the regression model. Likewise, due to the inclusion of fixed effects estimators, Stata dropped all players who have never quit a tournament (again, perfect prediction). All players who competed in fewer than 50 matches were also dropped from the dataset. This was due to calculation problems with Stata; when players with fewer than 50 matches played were included in the analysis, the logit regression would not compute an output either due to problems with the program's computational power or due to the fact that including these observations could force the logit regression into an "infinite loop" in which a maximum log likelihood value was never reached. This resulted in the final combined dataset of 29,547 observations.

All Non-Grand Slam tournaments were dropped to create a separate Grand Slam dataset. Because no retirements occurred in the final two rounds of these tournaments, all semifinal and final round observations were dropped. This resulted in the final Grand Slam dataset of 6,185 observations.

It was then necessary to ensure that all variables involving prize money could be compared on a standardized basis. All variables involving prize money in a currency other than USD were converted from their origin country currencies to USD.[32] All currency values were obtained from the currency exchange website OFX (n.d.). Conversions were executed using the rate from the closest recorded date to the tournament's start date. Likewise, all values were adjusted for inflation by calculating the 2017 dollar equivalent values using the CPI inflation calculator tool from the US Bureau of labor statistics (n.d.).

After importing the data into Stata for analysis, all non-numeric columns (Tournament, Surface, Winner, Loser) were recoded into numbers using the "encode" function so that these values could be controlled for as fixed effects estimators. All information in the utilized dataset was obtained, scraped, and recoded through this format.

[32] Tournaments that did not record their prize money distributions in US Dollars recorded their information in either Euros, Australian Dollars or Great British Pounds.