THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF ECONOMICS

THE EFFICIENCY OF PRIZE MONEY STRUCTURE OF WTA TOURNAMENTS

YUAN WANG
FALL 2013

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Economics and Mathematics
with honors in Economics

Reviewed and approved* by the following:

David Shapiro
Professor of Economics
Thesis Supervisor

Russell Chuderewicz
Senior Lecturer of Economics
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

In this thesis, the previous and current tournament profiles on the WTA Tour will be studied based on the assumptions of the Tournament Model, which was presented in Rosen (1986). In addition, a series of statistical tests are conducted to test the effects of prize money spreads and prize money breakdown of a tournament on the performance of participating players. The tests use the latest data available online from 2007 to 2012, and the data set contains statistics of tournaments of all levels. Equally interesting, studies of the Tournament Theory so far almost focus on ATP Tour of men, while this thesis instead chooses a new perspective from WTA Tour of women, therefore the results surprisingly show significant difference from the results of previous works.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First of all, I would like to show my appreciation to Dr. David Shapiro for being my thesis supervisor and nicely doing this job. His profound knowledge in economics and helpful suggestions inspire me beyond words. In addition, I would like to thank my fellow scholars who also participate in the honors thesis seminar for providing insightful comments on my thesis. Last but not least, I would like to say thanks to my lovely parents who give me my life and always support me no matter what I choose to do. My thesis cannot be completed without their efforts.

Chapter 1

Introduction

In recent years, the WTA (Women's Tennis Association) put considerable efforts in promoting women's professional tennis around the world. One part of the promotion is the replacement of its previous tournament profile with a new one that has a significant increase in total prize money of all tournaments on the tour. However, an increase in total prize money is hardly the guarantee to succeed in promoting the tour for the following reasons. First of all, hundreds of players are competing to earn a share of the prize purses and the earnings are of great significance to them because most of the players depend on the prize money to make ends meet. Equally important, in order to maximize their profits, tournament organizers hope to see participating players present quality performance to increase the ticket selling and to attract more TV contracts (Ivankovic, 1995). Last but not the least, some people argue that the WTA does not deserve much in the way of resources since women's tennis is less intense and less competitive than men's tennis. Due to the above reasons, it is interesting to study the prize money distribution methods in a tournament and try to seek the best prize money breakdown to have players fully respond.

In my thesis, I am going to test the previous and current tournament profiles on the WTA Tour based on the assumptions of the Tournament Model, which was presented in Rosen (1986). In addition, a series of statistical tests are conducted to test the effects of prize money spreads and prize money breakdown of a tournament on the performance of participating players. The tests use archival sport data collected from online resources, including the official site of the WTA and

some major tennis bet websites, which are free to use.

Chapter 2

Theoretical Frame an d Literature Review

In professional sports, there are a couple of factors that affect the players' performance: monetary reward is the prominent one, and in addition, the points that the ranking system is built on and the heterogeneity in the quality of contestants also play important roles. Besides, some external factors like tournament surface (hard, clay or grass), weather, courtesies of spectators and so forth may contribute to the efforts of competitors as well.

Section I. Rosen and His Tournament Theory

In "Rank-Order Tournaments as Optimum Labor Contracts" Lazear and Rosen (1981) suggest that the reward scheme based on workers' relative positions in an organization can produce the same incentive structure for the risk-neutral workers as does the marginal product. In addition, to reward workers according to their relative positions is less costly than to observe and measure each individual's exact output directly. However, the professional women's tennis tournaments adopt a special framework – single-elimination format, which means a player is automatically eliminated from a tournament after losing a match (WTA 2013 Official Rulebook). In this scenario, players cannot be regarded as workers who constantly make contributions to a single firm. Rosen in his "Prizes and Incentives in Elimination Tournaments" (1986) provides a new model that examines the efficiency of the prize money structure of a single-elimination tournament. This model also fits tournaments with uneven contests (Lallemand et al. 2005).

Rosen (1986) explains that an extra weight on the top rank prize is necessary to maintain the incentives of surviving players to aspire to higher goals regardless of past performances. The author also interprets competing in an elimination tournament by analogy with climbing a ladder – competitors have quite a few rungs ahead to encourage them to move upward. However, fewer rungs are left along the way as competitors advance. Once they reach a certain point, competitors have to examine their marginal cost and marginal payoff. If the top prize is not large enough, the competitors who have achieved high ranks may rest on their laurels and slack off their attempts to climb higher. Hence, to set efficient top prizes and make competitors feel motivated to exert extra efforts is of great significance.

Section II. Other existing Literature

Ivankovic (1995) conducts three empirical tests on the professional men's tennis tour (ATP Tour) according to the Tournament Model that Rosen proposes. The hypotheses include the pattern of the prize money breakdown structure in a tournament and the performance of participants regarding a series of factors. In his paper, Ivankovic expects to see the spread (to be utilized and explained in the data set section of this thesis) increase at a linear or increasing rate as the rounds left to be played decrease and have a jump in the final round. Another objective is to see if the top four ranks (the winner, finalist and two semi-finalists) are awarded more than half of the total prize money purse. Concerning performance, the hypothesis predicts that the effort level of competitors would change as the spread level changes. The results of the first two tests are contrary to expectations. In tournaments at all levels in the ATP Tour, prize money spreads increase at a decreasing rate as the rounds left to be played decrease and a clear plunge is seen in the final round, with very few exceptions. As to the percentage of reward to the top four ranks, the result ranges from 31.95 percent to 41.17 percent, with the minimum seen in Grand Slam

tournaments (128 Draws) and the maximum seen in 53-Draw tournaments. Therefore, Ivankovic concludes that ATP tournaments of any level do not to set the prize money structure according to Rosen's Tournament Theory. On the other hand, the third test result complies with the hypothesis: the effort level, in terms of total points played and total games played, increases as the spread increases. This result is consistent when utilizing other sports data to measure the effort level, such as total breakpoints earned, total breaks converted and total aces played. Hence, Ivankovic suggests that though poorly structured, spread levels in ATP tournaments still cause direct changes in players' performance. In a word, the higher the spread becomes, the harder competitors play on court.

Salge (2010) is the most recent work on this topic. It is built on Ivankovic (1995) with some further development. Salge redoes the tests of three hypotheses that were examined in Ivankovic (1995) with updated data of 129 professional men's tennis tournaments played on the ATP Tour during 2008 and 2009. Salge draws similar conclusions to those of Ivankovic (1995), which indicate that in general the prize structures of ATP tournaments are not designed in correspondence to Rosen's Tournament Model but the increasing spread has a positive influence on players' effort level on court. Besides testing three main hypotheses that have been examined in Ivankovic (1995), Salge also tests another four hypotheses, and they are: (1) the chance of winning a match increases when a player wins the first set; (2) upsets (the relatively lower ranked player winning a match against his higher ranked opponent) are more likely to occur in ATP tournaments than in Grand Slam tournaments; (3) the size of the total prize money purse of a tournament does not affect the performance of the players but the quality of the draw; (4) less competition is seen in a highly uneven match. All four hypotheses that Salge constructs are accepted after careful empirical studies. Of the four hypotheses, the second one is specific to the scenario of men's professional tennis tournaments because all women's tournaments of any kind adopt a best-two-of-three-sets format.

Lallemand et al. (2005) is one of the very few existing works that focus on the women's professional tennis tour (WTA Tour) to study the Tournament Model. In this paper, two factors that may affect players' performance are carefully studied: (1) the level of prizes and prize spread depending on the absolute and/or relative performance of players on court and (2) the heterogeneity in abilities of players in draws. However, three scholars allow the assumption that women players may react differently to prize incentives. For instance, women are supposed to be underperforming in competition since they may believe that their extra benefit is not big enough to compensate for their a priori lack of abilities (Gneezy et al., 2004).

In their paper, Lallemand et al. (2005) focus on the final two rounds of all tournaments, say two semi-final matches and a final match instead of the big picture of all matches. Due to this reason, the prize spread is measured in a way similar to but more complicated than the way defined by Rosen (1986); the prize spread is the sum of prize money breakdown for winning the semi-final and the possible value of the future prize (prize money earned by winning the final match). However, the possible value of future prize depends on two factors: one is the ability of the player and the other is the ability of the potential opponent to compete with in the final match. In a word, the prize spread is weighted by the chance of advancing, not simply by presuming a half-half chance of advancing to both players and multiplying the prize money breakdown by .5. However, the way to calculate the chance of advancing is not described in detail; hence the accuracy is not available to estimate. The effort level or the performance of an individual player is represented by games won in a match. The heterogeneity is measured by the differential between the rankings of the paired players in a match, so the favorites (higher ranked players) face negative heterogeneity while the underdogs (lower ranked players) face positive heterogeneity.

The empirical findings support the positive and significant relationship between prize money spread and the players' performance, as suggested by the Tournament Theory. In addition,

the results confirm the hypothesis that in uneven matches, the favorites win more games while the underdogs become less performing. However, I have concerns about the results. I wonder if it causes censoring problems as the samples selected are not reflecting the whole situation of tournaments and some information will be lost when analyzing censored data (Leung et al. 1997): first of all, focusing on the final two rounds of all tournaments, in which the four players who are surviving usually are close in ability, the research excludes many observations of uneven matches, so the results may be biased towards the prize incentive factor. For example, in a semi-final match between the top seed and the fourth seed, they both may react positively to prize spread increase regardless of the tiny difference in their abilities and rankings. Since the research targets semi-final and final matches only, the scenario in the example above is not unusual. Secondly, in every single match, regardless of the level of the tournament, a player has to win at least 12 games to win a match: a player can prevail over her opponent by 6 – 1 6 – 0 or she can edge her opponent by 6 – 4 6 – 4. Due to the difference in the number of games the winners lose (1 to 8), it is intuitively appropriate to expect that the player who dominates the match (with 6 -1 6 – 0) shows greater effort than the other winner, especially in the final rounds where players are similar in abilities and have almost equal possibilities to win. However, the number of games won fails to demonstrate the great effort, because in both matches the winners win 12 games. Even though the results of Lallemand et al. (2005) are consistent with those of previous works and statistically significant by its own, I still wonder about its viability due to the two reasons I mention above. However, by pointing out the presumption that women may react differently from male players in professional tennis tournaments and incorporating both prize incentive and heterogeneity in a regression, this paper still plays an important role in building up the theoretical frame and methodology in my thesis.

Chapter 3

Data Set

Section I. Tournament Profile of the WTA (The Women Tennis Association) Tour

The tournament profile of the WTA Tour went through a revolutionary transformation in 2009 as the WTA made some major adjustments to its previous tournament structure and has been following the system ever since: Premier 5 of the new profile is roughly equivalent to Tier I of the old profile and Premier is similar to Tier II. International Tournaments replaced Tier III and Tier IV tournaments. Besides, four tournaments that were in Tier I and Tier II, including Indian Wells, Sony Open, Madrid Open and China Open, were upgraded and categorized into Premier Mandatory, a new level that all players currently ranked top 10 have the commitment to participate in. In addition, the total prize money purses for each level may differ in each year. A tournament is also apt to be promoted into a higher-level category if its total prize money purse is raised substantially and satisfies the requirement for a higher-level event. For example, Brisbane International has been upgraded from International to Premier since 2011, when its total prize money was increased from $220K to $650K, and consequently the participant profiles and the draw size were changed as well.

Below is a table briefly presenting the previous and current WTA annual tournament profiles.

Table 3-1: WTA Tournament Profiles before and after 2009

WTA Tournament Profile (Before 2009)			WTA Tournament Profile (Since 2009)		
Level	Quantity	Total Prize Money	Level	Quantity	Total Prize Money
GrandSlam[1]	4	<$ 23M	Grand Slam	4	> $ 30M
			Premier Mandatory	4	$ 4,800K to $ 5,900K
Tier I	9	$ 1,340K to $ 2,500K	Premier 5	5	$ 2,000K to $ 2,168K
Tier II	14	$ 600K	Premier	12	$ 634K to $ 740K
Tier III	17	$ 175K to $ 200K	International	29-32	$ 220K
Tier IV	13	$ 145K			

Source: the WTA official website.

From the table above, the entries under "Quantity" are subject to change throughout the years, except for the quantity of Grand Slams that has a tradition of being 4 for over a century. Premier Mandatory, which was first introduced in 2009, maintains 4 and is fixed within a foreseeable future. For the other categories, it is clear that since 2009, WTA retrenches the quantities of Tier I and Tier II tournaments - nine Tier I tournaments have been reduced to five (Premier 5 tournaments) and fourteen Tier II tournaments to twelve (Premier tournaments). A new level of four mandatory tournaments has been added between Grand Slams and Premier 5. The quantities of low-level tournaments maintain steady. To guarantee frequent media exposure all year round and maintain a wide population of women's professional tennis across the globe, WTA expands the total prize money purses of all tournaments, especially the Premier levels, in

[1] Four of the ITF's leading National Associations own and organize the Grand Slams, not the ATP tour or WTA Tour. In addition, the total prize money for Grand Slams varies among four tournaments and along the time.

order to attract top players into these tournaments. Hence, the premier mandatories have draws of the parallel quality to Grand Slam Tournaments in terms of seeding. In fact, the top 20 ranking players have to fulfill the commitment to attend the four premier mandatory tournaments to avoid the punishment of fines and points deduction. In addition, WTA redistributes the bids for hosting International tournaments across the world to promote this sport in some area where tennis is less popular. Those International tournaments constrain the numbers of top-ranking participants (usually set as at most 1 from the top 20 and 4 from the top 50) in order to create more chances for the lower-ranking players to seek breakthroughs in their careers and make ends meet.

Section II. Data Set

The data set is divided into two subsets; one is called Tournament and Performance Data Set and the other Prize Money Data Set. The division is inspired by Ivankovic (1995). However, in that paper, the data set is generated in accordance with ATP Tour (Association of Tennis Professionals for men), which has some significant differences compared to WTA Tour, so I make some appropriate adjustments: for example, I do not incorporate the variable SETS since women's tournaments of all levels play best of three setters while in men's game, the Grand Slam Tournaments and finals of some high level tournaments, like Olympics, play the best of five sets. In addition, some trivial variables in the data set of Ivankovic (1995) are also dismissed because they are of little use to the analysis in this thesis, such as variables YEAR, TOURNAMENT of which the mean and standard deviation are obtained but contribute little to the test of hypotheses.

The Tournament and Performance Data Set supplies the information of the annual tournament profile of the WTA Tour and the records of the performance of participants on the tour. Data consists of 5927 professional tennis matches played in 2007, 2008, 2011 and 2012, of which the tournaments range from International Tournaments to the greatest Grand Slam

Tournaments. Tournaments differ in draw size and numbers of participants. Currently there are 3 different draw sizes - 32 draws, 64 draws and 128 draws. However, for the purpose of encouraging top-ranking players to participate in the events, quite a few of the premier level tournaments offer the seeded players privileges of advancing directly into the second round without competing in the first round and such priority is called a Bye. Hence, based on three draw-types mentioned above, there exist some variants: 28 draws, 30 draws, 56 draws and 96 draws. In addition, there are two Year-ending Championship Tournaments (the second one was introduced in the WTA Tour in 2009) exclusively for 8 players who are qualified according to the performance throughout the whole season. But these two tournaments will be excluded in my data set because they adopt Round Robin instead of the usual Single Elimination format. In total, after adjustment (the criteria are mentioned later) I have the data of 1231 matches in the year of 2007, 1453 in 2008, 1602 in 2011 and 1664 matches for 2012.

The variables included in this set are: **WRANK** – the instant ranking of the winner of a match, **LRANK** –the instant ranking of the defeated player in a match, **RNDF** equals |**WRANK – LRANK**| and represents the difference in talents between two paired players in a match**, W/L** – valued 1 if the WRANK < LRANK and 0 vice versa, **RET** – retirement or withdraw and it is a binary variable as well. To measure effort levels players exert on court, TIME representing the actual time of a match can be utilized because longer matches are associated with higher effort level (Ivankovic, 1995). However, due to lack of records of time duration in general, we have to seek other measurements for the effort level and performance from players. There are a few that can be considered: **SETS** – the sets played in a match, usually turns out being 2 or 3 if no withdraw or retirement occurs. It is not a good indicator of effort level because a two-setter with two tiebreaks (7-6, 7-6) is not necessarily less intense than a three setter ending with 6-2, 3-6, 1-6. **UPSET** – an upset occurs when the lower-ranking player wins the match. It is a binary variable and equals the absolute value |**W/L -1**|. It reflects the challenges facing top-ranking players from

bottom-ranking players. **TLPTS** – total point played in a match. It is one of the most appropriate replacements for **TIME**. **TLBRK** – total break points obtained by players and it is also a plausible alternative indicator of the effort level. In women's professional tennis matches, the serves are usually generated in a speed above 100MPH, so the player who holds the service game is usually considered having a kind of advantage over the receiver. But if the receiver plays aggressively enough and earns multiple break points, it shows a high effort level from players. If a match is close enough, both players may have quite a few chances to break their opponent's service games. For example, in one of the semifinals of the 2010 Miami Open, 25 break points were obtained during the 32 service games played. That match lasted for 155 minutes. **ACES** – the points won when the receiver fails to touch a good ball served by her opponent. In this paper, I am apt to mainly use **TLPTS** and **BRKPTS** to measure the effort levels.

The second data set, Prize Money Data Set, provides the variables that show the prize incentives of a match to players. In this data set, I generate a few variables especially for further analysis: **BRKDWN** – the prize money breakdown, which is used to indicate the difference between the prizes awarded to two sequent rounds. In another word, it is the marginal benefit of a player's advancement to the next round. $BRKDWN_i = P_{1+1} - P_i$, where P_i means the prize money compensation to the players eliminated in i^{th} round and in turn P_8 represents the prize money awarded to the champion in a 128-draw event. **SPREAD** – the marginal *possible* benefits of advancing into the next round. SPREAD deserves more explanation. According to Rosen (1986), SPREAD is defined as the guaranteed prize money for advancing into the next round plus the discounted sum of successive interrank rewards that could be achieved in the future:

$$SPREAD_i = BRKDWN_i + \beta_{i+1} * BRKDWN_{i+1} + (\beta_{i+2})^2 * BRKDWN_{i+2} + (\beta_{i+3})^3 * BRKDWN_{i+3} + \dots$$

, where β_i stands for the probability of advancing. For example, in a 128 Draw Grand Slam Tournament with 7 different rounds, there are 7 BRKDWN's (i=1, 2, ..., 7). Assume a player has survived into the third round in a 128-draw

event, and then she has 5 more possible matches to play in order to win the title. Hence, $SPREAD_3 = BRKDWN_3 + \beta_4 * BRKDWN_4 + (\beta_5)^2 * BRKDWN_5 + (\beta_6)^3 * BRKDWN_6 + (\beta_7)^4 * BRKDWN_7$. Rosen assumes that β is a constant with a value of ½ in his model, which means that for all matches, both players have a 50 percent chance to win. The assumption may be unrealistic in some matches, for example, where the top seed competes against a qualifier, who is ranked beyond 100, but there are only 32 seeds in a tournament and not even every seed has the absolute power to dominate matches, so approximately matches are played between players who have an equal probability to win. In addition, as the tournament progresses, the players surviving in the draw are close in abilities, so the 50 percent chance of winning a match for both players is reasonable, regardless of the minor difference in players' rankings. In terms of the levels of the tournaments, we have a few measures: **TLPRZ** – the purse of total prize money of a tournament; usually the larger the value is the higher level the tournament is at. **RKWGT** – the weighted average of instant ranks of all participants in a tournament, and it is an indicator of the quality of the draws and the overall ability of the participants. In order to incorporate all the factors that affect the effort level, a few other variables are necessarily to be taken into consideration: **SURFACE** – types of surfaces on which the tournaments are played and **SURFACE** contains three binary variables – **HARD**, **CLAY** and **GRASS**, representing the three main types of surfaces of WTA Tour tournaments, hard, clay and grass. There are subdivisions in each surface type category, for example within hard court surface, Australian Open is played on Plexicushion hard courts and US Open is played on acrylic hard courts, but in general the difference is trivial so that can be ignored. Due to the nature of the different surfaces, the balls move in different speed on different surfaces. It is commonly regarded that the ball traverses faster on the grass surface and relatively slowly on the clay surface. **RKDF**- the absolute value of the difference of instant rankings of the two players in a match, and it also has an effect on **TLPTS** or **BRKPTS**

since two players with similar rankings have similar abilities to play, which may lead to a longer and closer match.

Chapter 4

Hypotheses

Tournament Theory was invented by Lazear and Rosen (1981) and Rosen (1986). The theory has been used to describe different situations where the wages are not based on the marginal product of workers but depend on the rank-order of individuals. The practice in women's professional tennis is similar to what the tournament theory suggests: the players who are eliminated in the same round in a tournament get the same payment, regardless of their performance during the match, and the tournament winner gets the highest payment. In this sense, women's professional tennis prompts players to win more matches rather than to try their best to make the match enjoyable to the viewers.

When applying tournament theory to the women's professional tennis tour, we expect to see an increasing effort level corresponding to an expanding prize money spread.

Based on Rosen (1986), Invakovic (1995) and Salge(2010), I have structured a few hypotheses that are interesting to study in the women's professional tennis tour with the newest sports data of the years 2007, 2008, 2011 and 2012. The data resources available to the public start from the year 2007 to 2012. I randomly pick these four years, 2 for each before and after 2009, in which the tournament profile changed.

1. In order to maintain the incentives of fair performances from surviving players, SPREAD is set to increase linearly or at an increasing rate as the tournament progresses into further rounds. In addition, the variable SPREAD has a significant jump at the final round.
2. The quality of draws is positively related to the total prize money purse. That is to say, RKWGT is getting smaller as the TLPRZ increases.

3. As long as the prize incentives (SPREAD AND BRKDWN) increase, the effort level from the players increases as well. The following two regressions can be run to test this hypothesis:

$$\text{Lg (TLPTS)} = \alpha_0 + \alpha_1 \text{ lg (SPREAD\_10K)} + \alpha_2 \text{ lg (BRKDWN\_10K)} + \alpha_3 \text{ lg (RNDF)} + \alpha_4 \text{ lg (FIELD)} + \alpha_5 \text{ HARD} + \alpha_6 \text{ CLAY} + \mathcal{E}$$

$$\text{lg (TLBRK)} = \gamma_0 + \gamma_1 \text{ lg (SPREAD\_10K)} + \gamma_2 \text{ lg (BRKDWN\_10K)} + \gamma_3 \text{ lg (RNDF)} + \gamma_4 \text{ lg(FIELD)} + \gamma_5 \text{ HARD} + \gamma_6 \text{ CLAY} + \delta$$

It is reasonable to expect the coefficients α_1, α_2, γ_1, γ_2 to be positive because higher prize incentives may encourage both players in a match to play harder; and the coefficients α_3, α_4, γ_3, γ_4 negative because if the differential in rankings is huge, lower-ranked player in a match may not try as hard as the high-ranked player, presuming she would lose anyway. The positivity of the coefficients α_5 and γ_5 are hard to guess, but the coefficients α_6 and γ_6 can be positive because it is commonly known that tennis balls traverse slower on clay courts then longer rallies are expected to see in a match played on clay courts.

Chapter 5

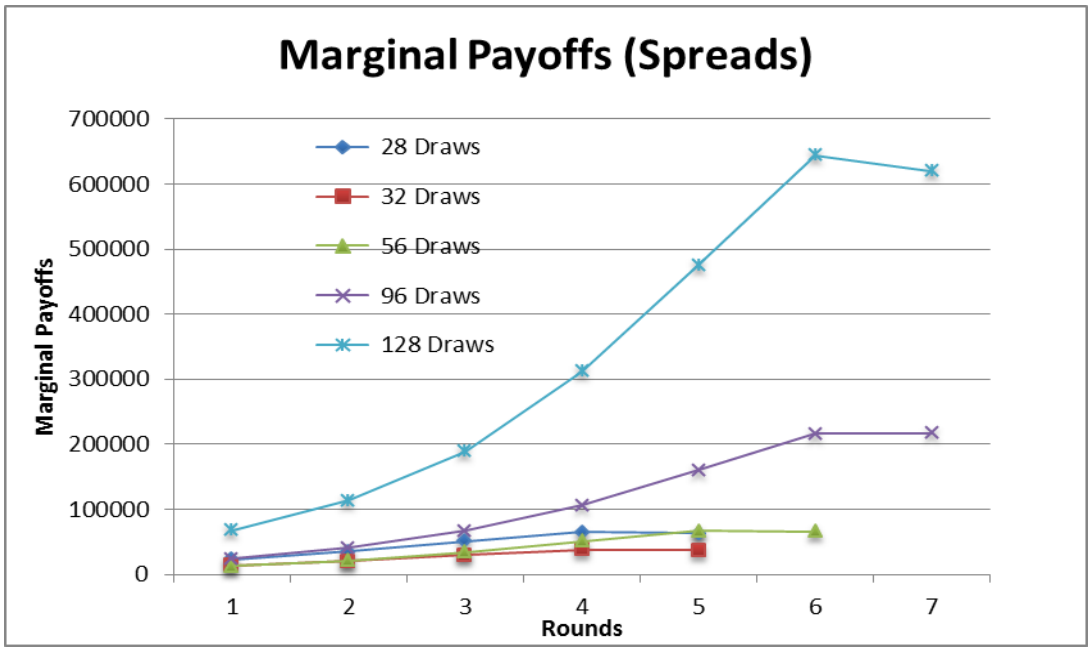
Results and Discussions

Section I. Test of Hypothesis 1

The study of hypothesis 1 is constructed with the variable $SPREAD_i$ as the dependent variable, where i represents the ordinal number of rounds in a tournament and i has the range between 1 and 7. We expect to see that in tournaments of all levels, as i becomes larger, $SPREAD_i$ goes up linearly or at an increasing rate. Besides SPREAD, I also incorporate the charts that show the prize money BREAKDOWN, where the vertical axis represents the extra money to earn by surviving to the next round. It is a more intuitive way to see the marginal benefits of winning one more match. Also we expect that the BREAKDOWN intends to increase dramatically as i increases.

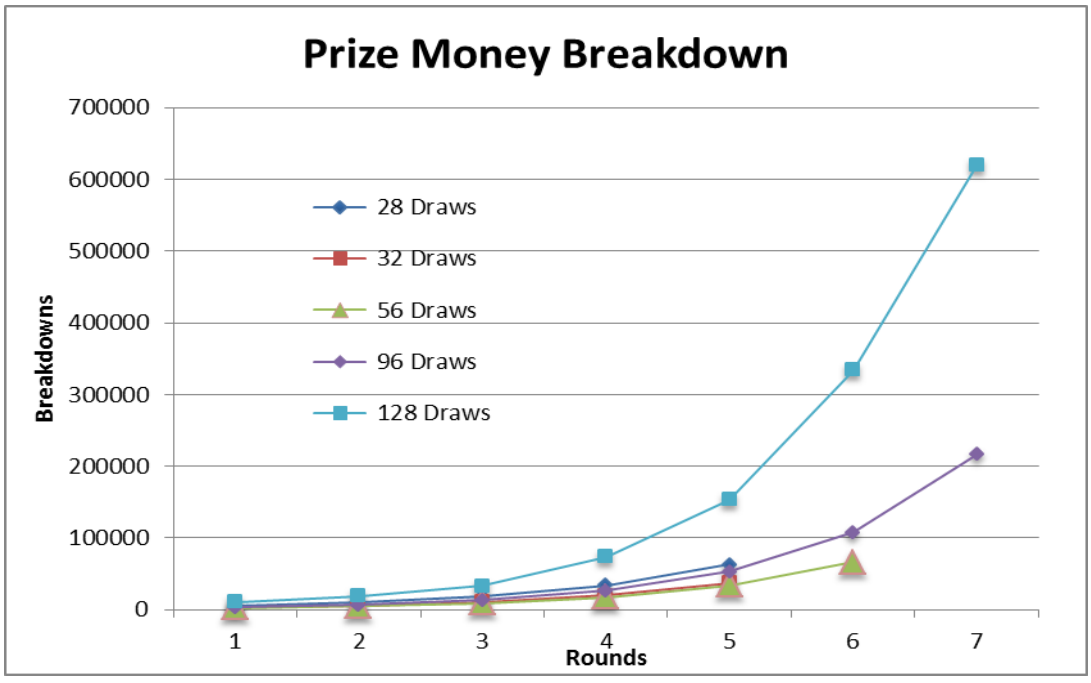
Below are the figures generated according to the data of prize money SPREADS and BREAKDOWNS collected from the WTA official site:

Figure 5-1: Marginal Payoffs for Draws 28, 32, 56, 96 and 128 before 2009



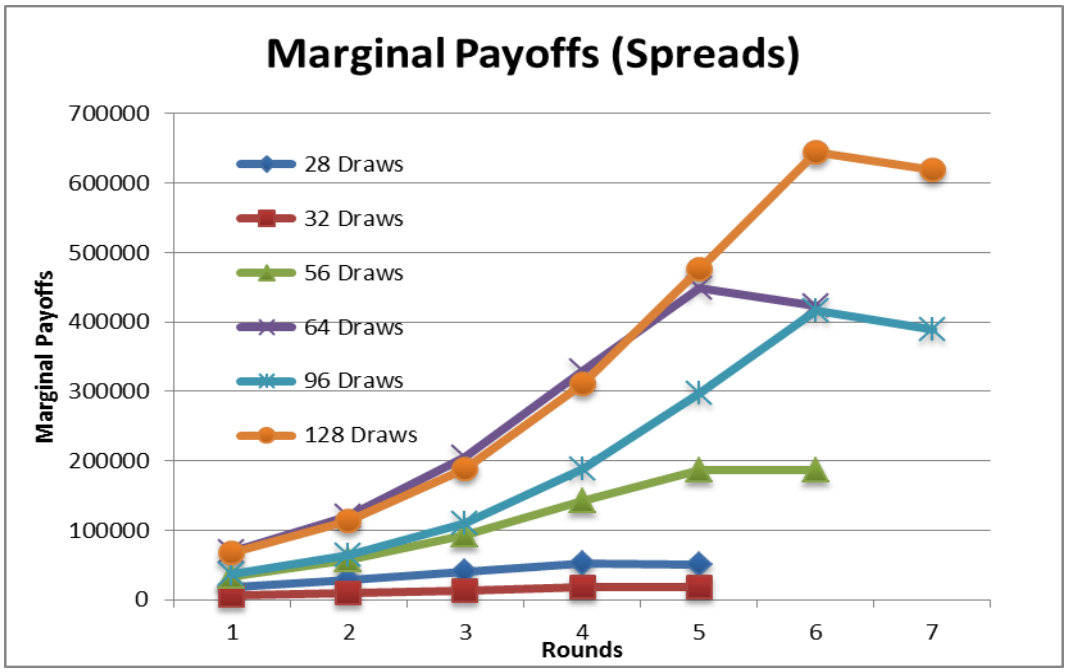
Source: the WTA official website.

Figure 5-2: Prize Money Breakdown for 28, 32, 56, 96 and 128 draws before 2009



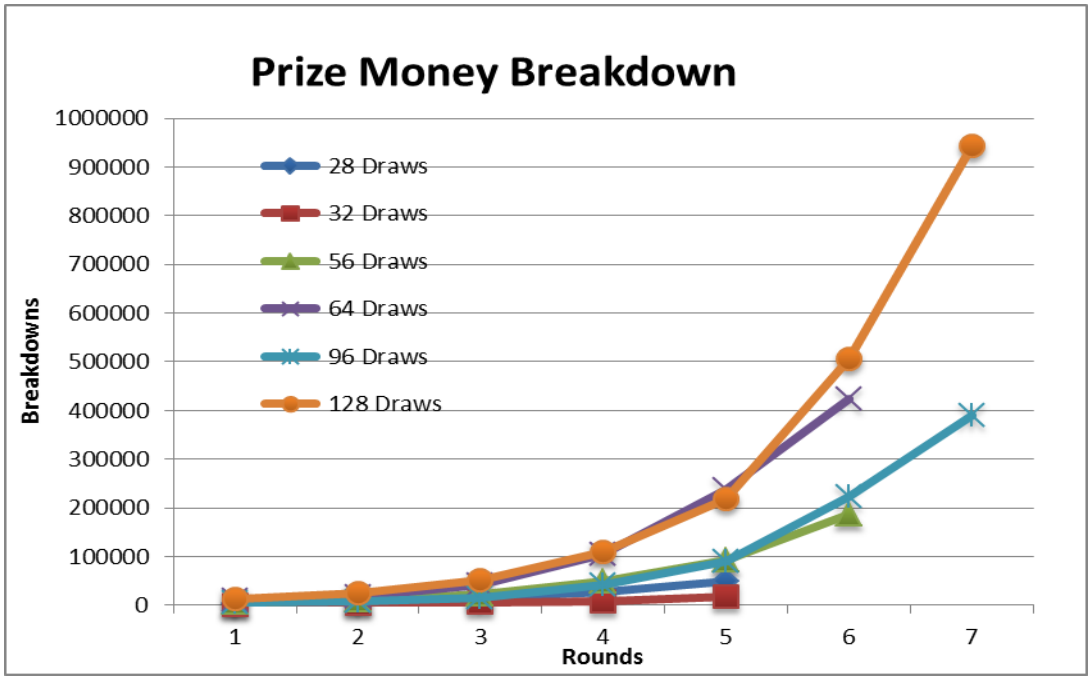
Source: the WTA official website.

Figure 5-3: Marginal Payoffs for Draws 28, 32, 56, 96 and 128 after 2009



Source: the WTA official website.

Figure 5-4: Prize Money Breakdown for Draws 28, 32, 56, 96 and 128 after 2009



Source: the WTA official website.

When studying Figure 5-1 and Figure 5-3, I conclude that Tournament Theory does not apply to the tournaments played in the WTA Tour, neither before 2009 nor after. The Grand Slam tournaments (128 draws) prize money distribution is not conducted according to Tournament Theory either. From Figure 5-1, it is clear to observe that before the final round, SPREAD lines incline upward, which indicates that before 2009 the WTA tournaments with all draw types have the marginal payoffs (spreads) increase at a relatively constant rate as the rounds left to be played decrease. However, the increasing trend disappears at the ending of all lines, and especially for that of 128 draw, a clear drop is observed. It shows that the marginal payoff (spread) for the final round of each draw type does not have a jump as the hypothesis predicts. It further indicates that the marginal payoff of winning the final match is not bigger than that of winning the semifinal in the previous round.

The same phenomena are also observed in Figure 5-3: Marginal Payoffs for Draws 28, 32, 56, 96 and 128 after 2009. Before the final match, the spreads of previous rounds increase as the players advance in a tournament. However, the marginal payoffs for winning the final match almost maintain the same as those of winning the semi-final match in 28-, 32- and 56- draw types, while in 64-, 96- and 128- draw types, the marginal payoffs for the final round experience a considerable drop. One noticeable thing about the lines of 96-draw and 128-draw tournaments is that SPREADS increase at an increasing rate except the last round, which is consistent with the prediction of the Tournament Theory.

Figure 5-2 and Figure 5-4 display another story of the increments of prize incentives as they incorporate prize money breakdown. Prize money breakdown can be obtained with the formula BRKDWN = $P_{i+1} - P_i$, where P_i means the prize compensation to the players eliminated in i^{th} round. It is obvious that the gaps go up at an increasing rate throughout all rounds, as long as the players survive in the tournament. The sum of the gaps from the first round to the round in which the player gets eliminated is the prize money that she earns from the tournament, and this

number is increasing dramatically if the player survives to the later rounds. It is also interesting to observe that in the tables in the appendix, the total prize money earned by the champion is approximately twice as much as the amount earned by the runner-up. This double relation also roughly stays solid between prize money earned by two consecutive ranks, for example, the prize money for a player eliminated in the third round is almost twice as much as that earned by a player losing the second round match.

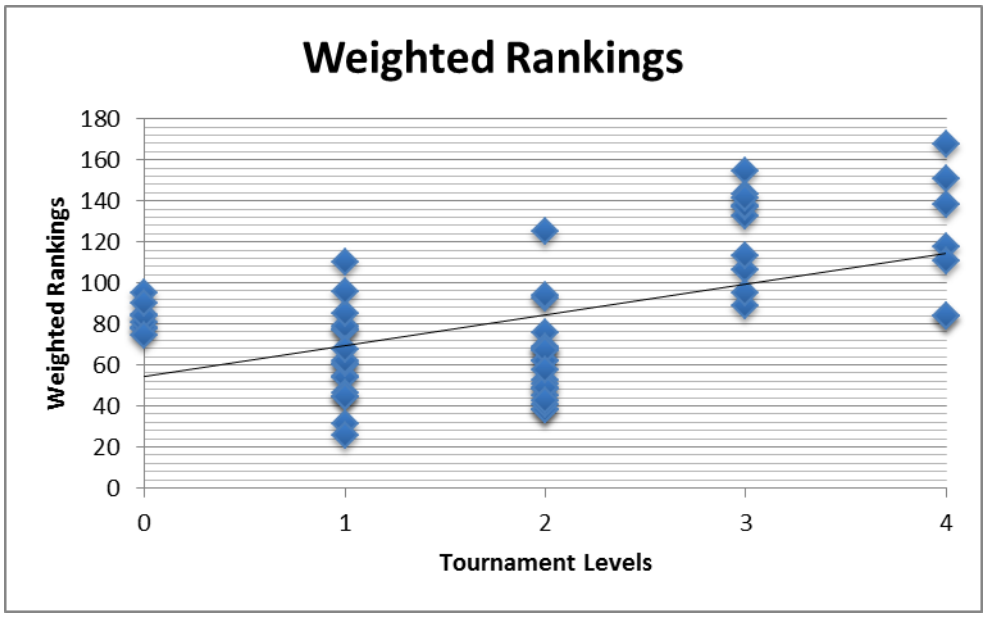
From the above graphs, we can conclude that the WTA tournaments are not conducted under the tournament theory. However, women's tennis is still regarded as one of the most competitive women's professional sport in the world. I can guess the reason to be that players do not necessarily do the math and calculate the spread; instead they base their decisions about what effort level to exert in a match on the prize money breakdown, of which the increment is appealing enough to motivate them to win an extra match – to pass through another round, the prize money almost doubles and that increase is specially large in 64, 96 and 128 draw tournaments. In addition, tennis being one of the most professionalized sports, professional tennis players have abundant endorsements and sponsorships from different corporations, and once players achieve great success in the WTA Tour, say ranking among the top five or winning a Grand Slam tournament, the income from endorsements would be several times more than the prize money they could earn from the tournament organizers. In recent years, a few young women were named the Highest-paid Athletes by Forbes mainly because they won a Grand Slam tournament or reach the top spot of the WTA Tour Ranking and attract attention from businesses. If we assume the commercial income to be a part of the benefits that come along with winning further rounds of a tournament, the potential benefits do significantly enlarge the marginal payoff (spread) of winning the final match.

Section III. Test of Hypothesis 2

According to Ivankovic (1995), the total prize money has a positive effect on tournament entry. Better and higher-ranked players are motivated to attend the tournaments that provide larger total prize money purses, and therefore, the weighted ranking of participants gets lower as the total prize money goes up, although no evidence indicates that the players are prompted to have better performances in these tournaments than tournaments with smaller total prize money purses. WTA tournaments are standardized tour-wise; tournaments of the same level have identical total prize money purses and distribution methods, therefore we consider the horizontal axis in the chart as the levels of tournaments. In the first chart, 0 stands for the Grand Slams tournaments, with total prize money of more than 8 million U.S. dollars; 1 represents the Tier I tournaments that provide $1,340,000 for all participants; 2 for Tier II with $600,000; 3 for Tier III with $175,000, and finally, 4 for Tier IV with $145,000. In the second chart, 0 represents the Grand Slam Tournaments; 1, the Premier Mandatory Tournaments; 2, the Premier 5 Tournaments; 3, the Premier Tournaments and 4, the International Tournaments, all with $220,000 U.S Dollars. Since weighted rankings are the arithmetic mean of instant rankings of all participants, the differential in sample size is no longer a problem.

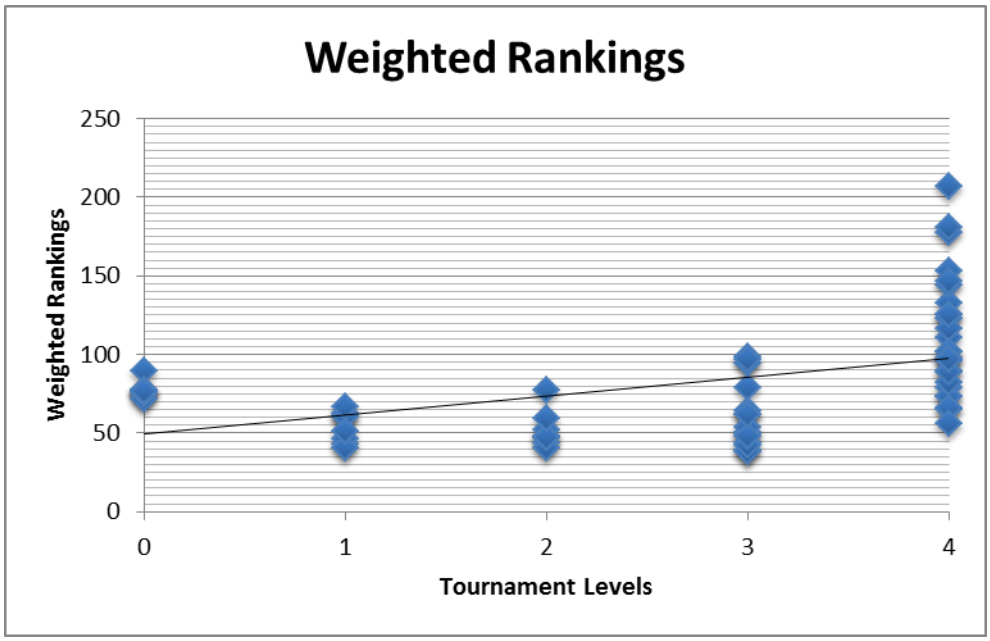
Below are two charts that display the relationship between total prize money purses and the weighted rankings of participants. In the chart, a straight trend line is drawn to help show the relationship.

Figure 5-5: Weighted Rankings of tournaments in 2007 and 2008



Source: the WTA official website.

Figure 5-6: Weighted Rankings of tournaments in 2011 and 2012



Source: the WTA official website.

In Figure 5-5, the positive and linear relationship is not as solid as presumed: the Grand Slam tournaments do not have the lowest Weighted Ranking among all tournaments and the Weighted Ranking of 1 approximately equals that of 2. However, we do observe a significant jump from 2 to 3, with the means around 60 and 110, respectively. In Figure 5-6, a similar observation can be made: 0 has higher weighted rankings than 1, and the weighted rankings are experiencing a constant increase among 1, 2, 3 and 4. I also rely on STATA to test the relationships between Total Prize Money and Weighted Rankings:

Table 5-1: Relationship between Weighted Rankings and Total Prize Money

Dependent Variable: Weighted Rankings (RKWEGT)		
Independent	w/ Grand	w/o Grand Slams
Total Prize Money	-.0010271	-.012816***
Constant	84.43186***	95.61796***
Observations (N)	127	111

Notes: *** indicates significance at the 1% level.

Therefore, we reject the hypothesis if we take the Grand Slam tournaments into consideration, but we accept the hypothesis if we exclude the Grand Slam tournaments and focus on tournaments managed by the WTA. One explanation for the deviation that the Grand Slam tournaments have bigger weighted rankings than do Tier I tournaments could be that fact that the Grand Slam might typically keep a certain number of slots and give a chance to compete in the tournament to some wild-card holders who are not qualified for the main draw considering their relatively low rankings. For example, in the 2007 Australian Open, of 128 draws 8 are provided to the local favorites, and the average ranking of the wild-card participants is 301.25, which is significantly larger than the weighted ranking of the whole tournament, 95.25. In addition, some qualifiers (having a ranking lower than a direct qualifying bottom-line to the main draw) begin to

compete from the qualifying rounds prior to the main tournament. Those qualifiers who pass through to the main draw also drag down the weighted ranking of the participating pool.

Section IV. Test of Hypothesis 3

In this section, the relationship between the prize money incentives and the paired players' actual performances will be tested. According to Rosen (1986), the value of maintaining fair performances from a player depends on the sum of sure prize money she would receive after advancing through the next round and the prospective prize money she could earn by winning more matches that come after. This monetary reward can be measured by SPREAD, the marginal payment of advancing defined by Rosen (1986) and BRKDWN, the extra prize money a player can earn by winning the current match. In addition, the outcome function is also affected by a couple of other factors: the ability of a specific player and the ability of the prospective opponent. It is expected that as the monetary rewards increase, the players give out better performance; the players may run faster and hit the ball harder. The mutual efforts in a single match make the game tight and it can be measured by the time a match lasts. However, the data for the time of multiple matches are not available (or recorded for public use), so I am seeking the alternative measurable variables – total points and the total break points of a match earned by both players.

In addition, in his paper Rosen (1986) also indicates that mismatch triggers less effort from players and especially the underdogs who are relatively lower in the WTA ranking system would regard their cost of continuation way higher than the favorites who are seeded and enjoy byes and other priorities in tournaments. Lallemand et al. (2005) suggest that according to the incentive hypothesis, in an uneven match the underdog may be less performing while the favorite will work relatively harder because she may see a higher chance of advancing and a higher value

of continuation. In this situation, total points played in a match would be expected to be decreasing as the degree of heterogeneity is increasing.

Concerning the performance on court, the following variables are utilized to conduct the tests: TLPTS, total points played, can be used to replace TIME because the data on time spent to finish a match is not available generally. Also the longer a match lasts, the more total points are to be played on court as it is acceptable that the total points of a match are positively and linearly related to the time of the match. Besides, each point can be earned within different time periods; for instance, an ace takes less than 10 seconds from the preparing bouncing to an actual strike while a 20-shot rally may take more than 2 minutes. Hence, using TLPTS instead of TIME may be more accurate in some aspect. TLBRK, total break points earned, for the same reason, can be used to measure the closeness of a match. From the beginning of a game, to earn a break point the receiver has to survive at least 3 high speed serves (to reach 0-40) generated by her opponent who is supposed to have the advantage over the receiver. If a breakpoint comes after a deuce (40 – 40), the pressure the server feels to maintain the advantages and the motivations the receiver holds to take the chance make the game more intense. If more breakpoints are achieved, no matter whether the breakpoint is converted or not, it is plausible that the players on the court are making great efforts to win the match.

To study the effect of monetary factors and heterogeneity on players' performance, I incorporate SPREAD (10k), BRKDWN (10K), RNKF, FIELD and two dummy variables HARD and CLAY as the independent variables and TLPTS as the dependent variable in my regression. After taking the natural logarithm of the independent variables SPREAD (10K), BRKDWN (10K), RNDF, FIELD and dependent variable TLPTS, I construct the following regression:

$$\lg(\text{TLPTS}) = \alpha_0 + \alpha_1 \lg(\text{SPREAD\_10K}) + \alpha_2 \lg(\text{BRKDWN\_10K}) + \alpha_3 \lg(\text{RNDF}) + \alpha_4 \lg(\text{FIELD}) + \alpha_5 \text{HARD} + \alpha_6 \text{CLAY} + \varepsilon$$

The variables have different scales and ranges. For example, TLPTS has the constraint within 69 to 271, while SPREAD (10K) takes its value between .1225 and 89.3569. In order to make the variables spread uniformly within similar ranges, I use natural logarithmic transformation to modify the data set and now log (TLPTS) has a range of 4.23 to 5.60 and log (SPREAD_10K) changes between -2.10 and 4.49.

Table 5-2: Effects of prize incentives and heterogeneity on women tennis players' performance in selected 35 tournaments in 2012

Source	SS	df	MS		
Model	2.61527979	6	.435879964	Number of obs =	1641
Residual	131.874379	1634	.080706474	F(6, 1634) =	5.40
				Prob > F =	0.0000
				R-squared =	0.0194
				Adj R-squared =	0.0158
Total	134.489659	1640	.08200589	Root MSE =	.28409

tlpts_log	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
spread10k_log	-.0086968	.0055509	-1.57	0.117	-.0195843	.0021908
brkdwn10k_log	-.0111673	.0105376	-1.06	0.289	-.031836	.0095014
rndf_log	-.0309146	.0070002	-4.42	0.000	-.044645	-.0171843
field_log	.0147741	.0159174	0.93	0.353	-.0164465	.0459948
hard	-.0083829	.0218272	-0.38	0.701	-.0511952	.0344293
clay	.0041109	.0244496	0.17	0.866	-.0438449	.0520667
_cons	4.968229	.0658082	75.50	0.000	4.839152	5.097306

Source: Author's calculation from data generated on WTA Tour website and tennis-data.co.uk

Table 5-2 blends all 1641 matches in 35 tournaments of all five levels that were held in 2012. Of 35 selected tournaments, 4 Grand Slam Tournaments and 4 Premier Mandatory Tournaments are all included, and in addition 4 Premier 5 tournaments, 8 Premier Tournaments and 15 International Tournaments that were selected at random. Any match that is not completed, in which either player chooses to retire or withdraw, is not included in the data set,

because some external factors such as injuries and personal issues, other than prize incentives and heterogeneity, may dominate the effort level that players exert on the court. Besides, in order to make the data set plausible, I modify it by eliminating the observations with the extreme values of spread10k. These observations consist of finals of 4 Grand Slam Tournaments and Premier Mandatory Tournaments and other matches, of which the winners, as mentioned before, can expect to receive a considerable amount of sponsorships and endorsements. This kind of potential financial benefit cannot be measured but certainly has a positive effect on professional players' reactions.

The two prize incentive factors, spread (10k) and breakdown (10k), and the ranking difference of two players in a match all have negative effects on the total points played in a match. However, sadly neither spread (10k) nor breakdown (10k) has a significant coefficient at the 10% level. I guess it is because spread (10k) and breakdown (10k) are highly correlated as the correlation [spread (10k) and breakdown (10k)] = .6612.

Table 5-3: Correlation between spread (10k) and breakdown (10k)

Independent	Spread	Breakdown
Spread (10K)	1.0000	
Breakdown (10K)	0.6612	1.0000
Observations N	1641	1641

Notes: *** indicates significance at the 1% level.

Hence, by using only one of these two variables at a time, we can study the effect that each of the two monetary factors, spread (10k) and breakdown (10k), has on total points earned in a match.

Two regressions are constructed:

$$\lg(TLPTS) = \beta_0 + \beta_1 \lg(SPREAD\_10K) + \beta_2 \lg(RNDF) + \beta_3 \lg(FIELD) + \beta_4 HARD + \beta_5 CLAY + \eta$$

$$\lg(TLPTS) = \mu_0 + \mu_1 \lg(BRKDWN\_10K) + \mu_2 \lg(RNDF) + \mu_3 \lg(FIELD) + \mu_4 HARD + \mu_5 CLAY + \delta$$

Table 5-4: Effects of prize incentives and heterogeneity on women tennis players' performance in selected 35 tournaments in 2012

Dependent Variable: Total Points Played [log (TLPTS)]

Dependent Variables	w./ spread (10k)	w./ brkdwn (10k)
log [SPREAD (10K)]	-.0115235***	
Log [BRKDWN (10K)]		-.0191009**
log (RNDF)	-.0311555***	-.0320071***
log (FIELD)	.0205132	.022307
SURFACE		
HARD	-.0072793	-.0049298
CLAY	.0050501	.005945
GRASS		
Constant	4.94622***	4.930936***
Observations N	1641	1641
Adj. R^2	0.0158	0.0150
F-stat	6.26	5.98

*Notes: ***/**/* indicate significance at 1/5/10% level, respectively.*
Source: Author's calculation from data generated on WTA Tour website and tennis-data.co.uk

The results presented in the above table reveal a negative effect of marginal payoff (Spread) on total points the paired players play in a match. For example: a 1 percent increase in the SPREAD (10K) will decrease the total points played by the paired players in a match by approximately .0115 percent. On the other hand, the gap between two players' talents also plays an important role. Measured by the difference in instant rankings of paired players, the gap increases when RNDF increases and this difference also has a negative influence on the total

points played in a match: if the difference in ranking increases by 1 percent, the total points played in a match decreases by .0312 percent. The potential opponents' rankings are also apt to have a positive influence on players' performance: an increase in rankings of potential opponents or players who are still surviving in the main draw leads more total points to be played in a match. This result regarding potential opponents' rankings is reasonable because an increase in FIELD can be regarded as relatively lower talents of the potential opponents and the players are apt to perform better because the cost of advancing is relatively lower. However, the variable FIELD is not significant at the 90% level, so the influence may not be obvious. Two dummy variables HARD and CLAY are not significant either.

Equally interesting, the other regression regarding the prize money breakdown displays similar results: prize money breakdown and difference in instant rankings of the players have negative effects in total points played and the coefficients of both variables, BRKDWN_10K and RNDF, are significant at the 95% and 99% levels, respectively. Meanwhile, the potential opponents' rankings and the dummy variables about the surface types are not significant at the 90% level.

Another way to measure the efforts of players on court is to use total breakpoints both players face in a match. Ivankovic (1995) explains that in professional tennis matches having a serve is an advantage. Hence, if both players fight hard we should expect to see more break points to be generated. Two similar regressions are constructed featuring two independent variables log (SPREAD_10K) and log (BRKDWN_10K), respectively, and log (TLBRK) as the dependent variable:

$$\lg(\text{TLBRK}) = \gamma_0 + \gamma_1 \lg(\text{SPREAD\_10K}) + \gamma_2 \lg(\text{RNDF}) + \gamma_3 \lg(\text{FIELD}) + \gamma_4 \text{HARD} + \gamma_5 \text{CLAY} + \delta$$

$$\lg(\text{TLBRK}) = \Phi_0 + \Phi_1 \lg(\text{BRKDWN\_10K}) + \Phi_2 \lg(\text{RNDF}) + \Phi_3 \lg(\text{FIELD}) + \Phi_4 \text{HARD} + \Phi_5 \text{CLAY} + \zeta$$

Table 5-5: Effects of prize incentives and heterogeneity on women tennis players' performance (total breakpoints earned) in selected 35 tournaments in 2012

Dependent Variable: Total Breakpoints [log (TLBRK)]

Dependent Variables	w./ spread (10k)	w./ brkdwn (10k)
log [SPREAD (10K)]	-.0031862	
Log [BRKDWN (10K)]		-.0054801
log (RNDF)	-.0352782***	-.0354915***
log (FIELD)	.0556781**	.059672**
SURFACE		
HARD		
CLAY	.0068944	.0064159
GRASS	-.1256535***	-.1273984***
Constant	2.643867***	2.641162***
Observations N	1641	1641
Adj. R^2	0.0158	0.0150
F-stat	6.26	5.98

*Notes: ***/**/* indicate significance at 1/5/10% level, respectively.*
Source: Author's calculation from data generated on WTA Tour website and tennis-data.co.uk

Table 5-5 shows the result of testing the effects of prize incentives and heterogeneity on women tennis players' performance in terms of total breakpoints generated in a single match. The results this time present a different scenario: prize incentives, either prize spread or prize breakdown, become less significant in affecting women players' abilities in earning breakpoints in a match, while the heterogeneity consistently has a negative and significant effect on the performance. Field, which represents the quality of all players remaining in draw, becomes significant at almost the 99% level.

To stretch the research to the years 2007, 2008, 2011 and 2012, two regressions are constructed:

$$\lg(TLPTS) = \alpha_0 + \alpha_1 \lg(SPREAD\_10K) + \alpha_2 \lg(BRKDWN\_10K) + \alpha_3 \lg(RNDF) + \alpha_4 \lg(FIELD) + \alpha_5 HARD + \alpha_6 CLAY + \varepsilon$$

$$\lg(TLBRK) = \gamma_0 + \gamma_1 \lg(SPREAD\_10K) + \gamma_2 \lg(BRKDWN\_10K) + \gamma_3 \lg(RNDF) + \gamma_4 \lg(FIELD) + \gamma_5 HARD + \gamma_6 CLAY + \delta$$

Results are shown in the following tables:

Table 5-6: Effects of prize incentives and heterogeneity on players' performance i

	Dependent Variable: Total Points Played (TLPTS)							
Independent Variables	2007		2008		2011		2012	
log[SREAD(10K)]	-.08404**		-.07362**		-.01334***		-.01152***	
	(.02639)		(.00345)		(.00412)		(.00487)	
log[BRKDWN(10K)]		.08159**		-.07035**		-.01315**		-.01910**
		(.02812)		(.00401)		(.00483)		(.00925)
Heterogeneity								
log (RNDF)	-.02613***	-.02224**	-.03346***	-.03312***	-.03547***	-.03354***	-.03116***	-.03201***
	(.00931)	(.00927)	(.00873)	(.00845)	(.00154)	(.00207)	(.00700)	(.00697)
log (FIELD)	.02861	-.00685	-.02357	-.02168	-.024234	-.02146	.02051	.02231
	(.02858)	(.02642)	(.03829)	(.03459)	(.02145)	(.02459)	(.01497)	(.015181)
Tournament Characters								
HARD	-.00101	-.00591	-.00432	-.00317	-.00632	-.00562	-.00728	-.00493
	(.02256)	(.02260)	(.02341)	(.02153)	(.02458)	(.02537)	(.02180)	(.02173)
CLAY	-.01201	-.00951	.01354	.01356	-.00348	-.01373	.00505	.00595
	(.02354)	(.02361)	(.02436)	(.01938)	(.02442)	(.02354)	(.02443)	(.02443)
GRASS								
Constant	4.99999***	4.99985***	4.99953***	4.99983***	4.96732***	4.96241***	4.94622***	4.93094***
	(.10936)	(.10977)	(.08342)	(.08345)	(.05835)	(.05824)	(.06245)	(.06138)
N	1231	1231	1453	1453	1602	1602	1641	1641
Adjusted R^2	.0088	0.0013	.0112	.0108	.0142	.0163	.0158	.0150
F-Statistics	2.81	1.33	5.34	4.91	6.02	5.84	6.26	5.98

*Notes: ***/**/* indicate significance at 1/5/10% level, respectively.*

Source: Author's calculation from data generated on WTA Tour website and tennis-data.co.uk

Table 5-7: Effects of prize incentives and heterogeneity on players' performance ii

	Dependent Variable: Total Breakpoints (TLBRK)							
Independent Variables	2007		2008		2011		2012	
Prize Incentives								
log [SREAD(10K)]	-.10116***		-.00462		-.00872		-.00319	
	(.03907)		(.00836)		(.00907)		(.00738)	
log [BRKDWN(10K)]		-.01515		-.00395		-.00915		-.00548
		(.01295)		(.01783)		(.00812)		(.01381)
Heterogeneity								
log (RNDF)	-.01946	-.01478	-.07621*	-.07538*	-.03194**	-.03147**	-.03528***	-.03549***
	(.01378)	(.01370)	(.04820)	(.04821)	(.01368)	(.01348)	(.01034)	(.01030)
log (FIELD)	.03474	-.00793	.09275	.09597	.05347	-.05079	.05568**	.05597**
	.04231	.03905	(.03354)	(.04223)	.02423	.02540	(.02226)	(.02248)
Tournament Characters								
HARD	.11810***	.11220***	.01032	.01278	.003118	.003112	.00347	.00315
	(.03340)	(.03347)	(.02236)	(.02946)	(.05334)	(.05332)	(.04724)	(.03668)
CLAY	.16383***	.16685***	.02947	-.04392	-.00638	.00668	.00689	.00642
	(.03485)	(.03491)	(.02574)	(.02463)	(.02348)	(.02349)	(.02482)	(.02474)
GRASS								
Constant	2.68739***	2.68729***	2.62423***	2.63421***	2.67984***	2.68845***	2.64387***	2.64116***
	(.16190)	(.16228)	(.08473)	(.08584)	(.09182)	(.09248)	(.08660)	(.08531)
N	1224	1224	1453	1453	1602	1602	1641	1641
Adjusted R^2	.0217	.0171	.0147	.0145	.0139	.0134	.0160	.0160
F-Statistics	5.51	5.25	5.89	5.96	6.07	6.08	6.32	6.32

*Notes: ***/**/* indicate significance at 1/5/10% level, respectively.*

Source: Author's calculation from data generated on WTA Tour website and tennis-data.co.uk

From Table 5-6, prize spread has a negative and significant effect on the total points played in a match in 2007, 2008, 2011 and 2012, while prize money breakdown has a positive effect of women's performance in 2007, which is opposite to the results in the other three years. Rank difference of the paired players in a single match remains with a negative effect on total points played and the coefficient displays significance in all four years. The ability of all players staying in the draw fails to exert a significant effect on total points and neither does the surfaces of the tournament.

On the other hand, Table 5-7 shows that TLBRK has poorer performance in reflecting the effort level of players on court than TLPTS does, because all the variables related to prize money incentives have insignificant coefficients with the only exception of SREAD in 2007. In addition, the positivity of a specific coefficient may be inconsistent in different years, for example, the tournament-character variable CLAY has coefficients of which the signs switch throughout years, even though it showcases significance at 1% level in 2007. Therefore, I intend to dismiss the use of TLBRK as the dependent variable.

Chapter 6

Summary

Even though a new tournament profile has replaced the previous one since 2009, WTA Tour tournaments are not conducted based on Rosen's Tournament Theory. In addition, from the data of the years 2007 and 2012, professional women tennis players show negative reactions to both prize money spread and heterogeneity in ability.

The result of women players' reactions to prize spread is inconsistent with the existing literature. Most of the previous works focus on men's professional tennis tournaments. In general, men may tend to be more aggressive under competition than women regardless of the differential between rankings (Knight, 2002) and male tennis players are less dissimilar in physical conditions. For example, heights of women tennis players range from 5'6 to 6'2 and show a noticeable variance among the women players while most male players exceed 6' and are below 6'2. Also, women show greater difference in physical power and strength than men do. Focusing on the women's professional tennis tour, Lallemand et al. (2005) draw the similar conclusion to works on the men's tennis tour in terms of positive reactions to prize incentives, but the coverage of their data set is over the final two rounds in tournaments of choice, which expels the huge discrepancy in abilities that is shown in early rounds of tournaments. My results based on sports data up to the year of 2012 are likely to give a big picture of the current situation of women's tennis tour: there is huge heterogeneity in abilities among professional women players and the discrepancy handicaps underdogs exerting extra efforts when facing stronger opponents. In other words, the prize money distribution is not sufficient to encourage low ranked players to outperform while it makes higher ranked players see higher value of continuation than the lower

ranked players. Hence in general, when the prize spread increases, total points played in a match decreases to some extent.

Appendix A

The Prize Money Distributions of the WTA Tour Tournaments

28 Draws 2007 & 2008				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	5987.5	1	5192.5	22645.3125
2	11180	47.5	9657.5	34905.625
3	20837.5	90	18032.5	50496.25
4	38870	160	33592.5	64927.5
5	72462.5	245	62670	62670
6	135132.5	352.5		
32 Draws 2007 & 2008				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	1080	1	822.5	3827.03125
2	1902.5	17.5	1622.5	6009.0625
3	3525	32.5	3097.5	8773.125
4	6622.5	57.5	5835	11351.25
5	12457.5	90	11032.5	11032.5
6	23490	127.5		
56 Draws 2007 & 2008				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	2395	1	2297.5	13061.09375
2	4692.5	27.5	4482.5	21527.1875
3	9175	47.5	8805	34089.375
4	17980	90	17160	50568.75
5	35140	160	33655	66817.5
6	68795	245	66325	66325
7	135120	352.5		
96 Draws				

2007 & 2008				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	3852.5	1	3422.5	23536.46484
2	7275	27.5	6942.5	40227.92969
3	14217.5	42.5	13372.5	66570.85938
4	27590	67.5	26113.75	106396.7188
5	53703.75	115	52711.25	160565.9375
6	106415	217.5	107373.75	215709.375
7	213788.75	337.5	216671.25	216671.25
8	430460	482.5		
128 Draws 2007 & 2008				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	17966.375	3	10666.625	67355.30273
2	28633	80	18696.5	113377.3555
3	47329.5	125	33491.875	189361.7109
4	80821.375	210	73866.125	311739.6719
5	154687.5	375	153708.75	475747.0938
6	308396.25	675	334214.125	644076.6875
7	642610.375	1050	619725.125	619725.125
8	1262335.5	1700		

28 Draws 2011 & 2012				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	5037.5	1	4118.75	18147.1875
2	9156.25	60	7631.25	28056.875
3	16787.5	120	14715	40851.25
4	31502.5	300	27297.5	52272.5
5	58800	320	49950	49950
6	108750	470		
32 Draws 2011 & 2012				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	1725	1	1225	5882.5
2	2950	30	2480	9315
3	5430	70	4770	13670
4	10200	130	8800	17800

5	19000	200	18000	18000
6	37000	280		
56 Draws 2011 & 2012				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	5725	1	5312.5	34031.25
2	11037.5	70	10612.5	57437.5
3	21650	125	22575	93650
4	44225	225	49025	142150
5	93250	395	93250	186250
6	186500	620	186000	186000
64 Draws 2011 & 2012				
1	11,740	5	9,455	69791.24102
2	21,196	80	17,807	120671.6841
3	39,003	140	40,777	205729.4939
4	79,780	250	105,092	329904.0765
5	184,872	450	238,126	449624.6837
6	422,998	700	422,998	422997.625
7	845,995	1000		
96 Draws Mandatory 2011 & 2012				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	7,377	5	4,735	36585
2	12,113	50	8330.25	63699.5
3	20,443	80	16119.75	110738.5
4	36,563	140	40537.5	189237.5
5	77,100	250	88900	297400
6	166,000	450	222000	417000
7	388,000	700	390,000	390000
8	778,000	1000		

128 Draws 2011 & 2012				
Rounds	Prize Money	Points	BRKDOWN	Mgl Payoff
1	17966.375	3	10666.625	67355.30273
2	28633	80	18696.5	113377.3555
3	47329.5	125	33491.875	189361.7109
4	80821.375	210	73866.125	311739.6719
5	154687.5	375	153708.75	475747.0938
6	308396.25	675	334214.125	644076.6875
7	642610.375	1050	619725.125	619725.125
8	1262335.5	1700		

Bibliography

Ivankovic, Miren. 1995. "The Tournament Model: An Empirical Investigation of the ATP Tour." Ph.D. thesis, Anderson University, Anderson, South Carolina.

Lazear, Edward and Rosen, Sherwin. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *The Journal of Political Economy*, Vol. 89, No. 5 (Oct., 1981), pp. 841-864.

Rosen, Sherwin. 1985. "Prizes And Incentives in Elimination Tournaments (Working paper No. 1668)." Retrieved from National Bureau of Economic Research.

Rosen, Sherwin and Sanderson, A. 2000. "Labor Markets In Professional Sports. (Working Paper No. w7573)" Retrieved from National Bureau of Economic Research: http://www.nber.org/papers/w7573

Salge, Carolina. 2010. "The ATP World Tour: How Do Prize Structure and Game Format Affect the Outcome of a Match?" Master Degree Thesis, Clemson University, Clemson, South Carolina.

Leung, Kwan-Moon, Elashoff, Robert M. and Afifi, Abdelmonem A. 1997. "Censoring Issues In Survival Analysis". *Public Health* 1997. 18:83–104.

Gneezy, and A. Rustichini. 2004. "Gender and Competition at a Young Age." *Economic American Review*, 94, pp. 377-81.

Starr, Cynthia. 1988. *We have come a long way: the story of women's tennis Billie Jean.King*. Rockefeller Center, New York City. McGraw-Hill Press.

Ware, Susan. 2011. *Game, set, match: Billie Jean King and the revolution in women's sports.* Chapel Hill, North Carolina. University of North Carolina Press.

Tennis Tournament Data Files. Tennis-Data.co.uk. http//www.tennis-data.co.uk

WTA, Inc. (2012). Women's Tennis Association 2012 Official Rulebook. FL, USA.

WTA Tournament Archive. WTA, Inc. FL, USA. http://www.wtatennis.com/tournament-archive

ACADEMIC VITA

Yuan Wang

yzw5134@gmail.com
84 Simmons Hall, University Park, PA 16802

EDUCATION

The Pennsylvania State University Park — **University Park, PA**
Schreyer Honors College, The College of The Liberal Arts — *Degrees Expected December 2013*
B.S. in Economics (Economics Departmental Honors Program), B.S. in Mathematics
Named to Dean's List for four semesters in a row.

The University of California, Berkeley — **Berkeley, CA**
UC Berkeley Summer Session — *May 2011 – August 2011*
Relevant Course-work: Linguistics, Multivariable Calculus and Art History.

RELEVANT EXPERIENCE

Bates White REU Program (Research Experiences for Undergraduates) — **University Park, PA**
Award Recipient and Researcher — *February 2012 – May 2012*
Worked closely with Professor Alexander Monge-Naranjo (Ph.D. from University of Chicago, 1999) for 125 hours and finished a three-page report for Bates White Economic Consulting Firm. Only five students are selected each semester.

EXREACURRICULAR ACTIVITIES

Penn State Service Trip Planning Team — **University Park, PA**
Trip Coordinator and Leader — *October 2011 –May 2013*
Planed and lead service trip to East Coast cities, like Washington D.C., Philadelphia and New York to serve individuals including homeless people, seniors, women and children.

State College World Church Office Iraqi Refugee ESL Program — **State College, PA**
ESL Teaching Assistant — *December 2011 – May 2012*
Taught English to refugees from Iraq and other Middle East countries to help them fulfill duties at Public Assistance offices in Pennsylvania.
Attended all the classes taught by fellow teachers and all the discussion sessions to improve the contents of the course.

YSOP (Youth Service Opportunity Program) Trip — **Washington D.C.**
Trip safety advisor — *November 2011*
Served homeless people, paralyzed people and seniors with their Thanksgiving food supply and did research with the Penn State Office of Student Activities and YSOP, Food and Friends and other organizations.

Penn State Global Office, Global Connection Volunteer — **University Park, PA**

- Passport on a Plate gala dinner "A Global Odyssey" — *January 30, 2011*
- YMCA State Swimming Championships — *April 2, 2011*
- State College Area Global Children Festival — *April 9, 2011*