I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12012

# Career Lotto: Labor Supply in Winner-Take-All Markets

Wayne Grove
Michael Jetter
Kerry L. Papps

DECEMBER 2018

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12012

# Career Lotto: Labor Supply in Winner-Take-All Markets

**Wayne Grove**
*Le Moyne College*

**Michael Jetter**
*University of Western Australia, IZA and CESifo*

**Kerry L. Papps**
*University of Bath and IZA*

DECEMBER 2018



**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## Career Lotto: Labor Supply in Winner-Take-All Markets*

Are people prone to selecting occupations with highly skewed income distributions despite minuscule chances of success? Assembling a comprehensive pool of potential teenage entrants into professional tennis (a typical winner-take-all market), we construct objective measures of relative ability and earnings projections. We find that prospective tennis professionals are attracted to right-skewed earnings distributions, independent of mean and variance. If skewness in prize money fell to zero, males would be 23% and females 5% less likely to continue pursuing a professional career, on average. Thus, winner-take-all labor markets appear to systematically encourage those with modest talents to pursue long-shot careers.

**JEL Classification:** J22, J24, J31, J44, L83

**Keywords:** winner-take-all markets, superstar markets, labor supply, human capital, gender differences, skewness preferences

**Corresponding author:**
Michael Jetter
University of Western Australia
35 Stirling Highway
Crawley 6009
Australia
E-mail: mjetter7@gmail.com

* We are grateful to Lowell Tayor, Dennis Wilson, session participants at the American Economic Association, the European Society of Labour Economists, the Midwest Economic Association/SOLE meeting, the Western Economic Association meetings, and the Royal Economic Society annual meetings, as well as attendees at the research seminars at Curtin University, LeMoyne College, and the University of Western Australia for constructive feedback, comments, and suggestions. For funding, we are thankful to BHP Billiton, Le Moyne College's Research & Development Committee and the Fallon Professorship. All remaining errors are our own.

*"The contempt of risk and the presumptuous hope of success are in no period of life more active than at the age at which young people choose their professions."*

Adam Smith (1776)

# 1 Introduction

"Never give up on your dreams" is the inspirational message uttered by medalists and award winners in the arts, entertainment, and sports, as well as by entrepreneurs who founded globally-dominant companies. However, advice about occupational choices from the victors in winner-take-all markets suffers from survivorship bias: young aspirants almost exclusively observe the extraordinary outcomes of the winners but not the effort and investments of *all* potential market entrants. For example, after having her Harry Potter manuscripts rejected by 12 publishing houses, J.K. Rowling became the first billionaire author – yet only one in 15,000 submitted fiction manuscripts gets published (Caves, 2000) and most authors earn less than the minimum wage (Gibson et al., 2015). Similarly, Jeff Bezos made Amazon the largest e-commerce company and became the world's richest person (Brynjolfsson and McAfee, 2014) – but less than 0.01 percent of consumer application developers can be considered a financial success (Guglielmo, 2014). In general, winner-take-all markets offer greater rewards to fewer superstars than ever due to market integration and reduced cost as a result of recent technical changes of digitalization, telecommunications and networks (Rosen, 1981; Brynjolfsson and McAfee, 2014). However, there is no empirical evidence on what the existence of these markets does to occupational choices.

Adam Smith's conjecture is that the astonishing fame and fortune of luminaries – like J.K. Rowling, Jeff Bezos, Beyoncé, and LeBron James – encourages *too many* entrants, causing a socially inefficient allocation of resources (Frank and Cook, 2010). In contrast, Rosen and Sanderson (2001) speculate that continuous feedback on one's performance causes potential entrants to switch to more realistic careers when the prospects of success in a superstar labor market become sufficiently unfavorable, thus doubting that aspirants act as "giddy risk lovers with unrealistic assessments of themselves". The ability to compete in such winner-take-all markets, though, only reveals itself by on-the-job talent discovery, such as the success of a computer application, the submission of a book manuscript, or selection for a symphony orchestra or professional sports team. Thus, contenders in the superstar selection process must make substantial pre-market investments to develop their abilities or products. Consequently, potential entrants to these markets face negative rates of return, on average, since mean earnings are dwarfed by the direct expenses and opportunity costs associated with pre-market development and market participation (e.g., see Barberis and Huang, 2008).

Then how can we explain people's propensity to choose unfair gambles and enter superstar markets? Portfolio theory predicts that risk averse people dislike the variance of an income distribution (Arrow, 1965) but like its skewness (Kimball, 1990). Indeed, people tend to accept lower expected payoffs and a higher variance in return for greater skewness in labor markets (Hartog and Vijverberg, 2007), entrepreneurship (Chen et al., 2018), savings (Gollier, 2001), insurance (Barseghyan et al., 2013; Collier et al., 2017), and financial investments (Brunnermeier and Parker, 2005; Boyer et al., 2009; Green and Hwang, 2012; O'Donoghue and Somerville, 2018). Ample evidence from behavioral economics suggests that individuals make decisions under uncertainty based on the attractiveness of outcomes, not just their underlying probabilities (Kahneman and Tversky, 2013; Dertwinkel-Kalt and Köster, 2017).

Beyond winner-take-all labor markets, right-skewed income distributions also appear to influence occupational choices, e.g., among US male college graduates (Flyer, 1997). Substantial variation exists in the level of earnings inequality across professions, even within narrowly-defined occupations. Figure 1 shows that the Gini coefficient within 3-digit occupations in the US varied from approximately 0.2 for postal clerks to 0.7 for athletes in 2010. Figure 1 also reveals that inequality within occupations increased in over 60 percent of these occupations since

1980. In many cases, occupations that require similar skills to enter and exhibit similar average earnings have vastly different levels of inequality. For example, while aerospace engineers and those in financial services sales occupations earn roughly the same on average, inequality is more than twice as high in the latter and this gap has widened over time. Thus, might rising income inequality within occupations cause inefficient labor market resource allocation, similar to what has been argued regarding speculative finance fields (e.g., see Philippon and Reshef, 2012, and Bolton et al., 2016)?

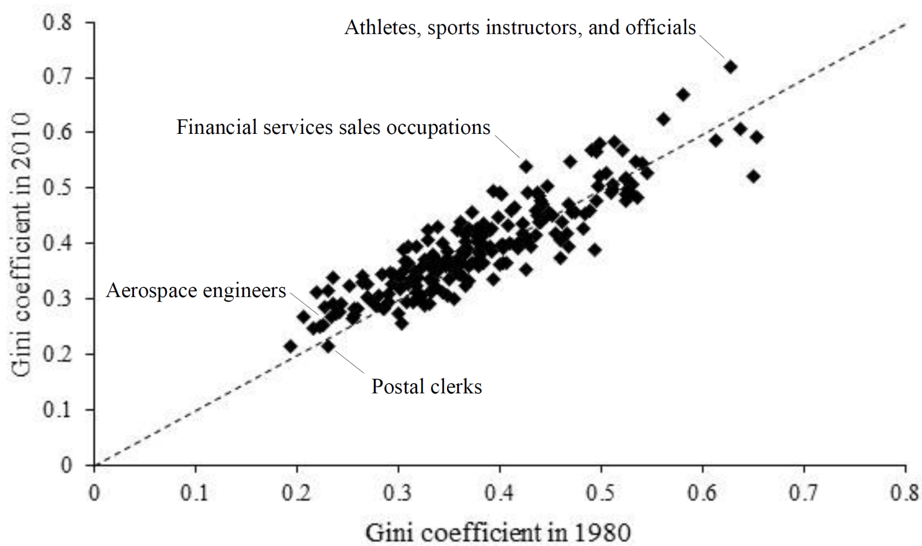


**Figure 1:** Gini coefficients within 3-digit occupations in the US in 1980 and 2010. Data on wage and salary income in the previous year from the 1980 Census and 2010 American Community Survey 1 percent samples are used. Only occupations with more than 500 workers in both samples are included.

The most basic dilemma for the families of young aspiring performers is that "data are not available to calculate meaningful success probabilities for potential entrants" (Rosen and Sanderson, 2001). Calculating such odds requires objective pre-professional and professional measures of ability for a comprehensive sample of potential market entrants and information about their likely earnings. The lack of such data explains why little empirical analysis exists of the success probabilities of young aspirants in winner-take-all markets, even though the notion that ability early in life predicts career outcomes underlies the studies of talent development (Baker et al., 2017), prodigies (Lubinski et al., 2014), expertise (Ericsson et al., 2018), and talent selection (Berri, 2005, p.360-373). More broadly, decision-makers under uncertainty cannot learn how market returns and risks are jointly distributed because they only observe idiosyncratic market signals about projects that have actually been implemented and those who have entered the market (Jehiel, 2018). As a consequence, such sample selection and survivorship bias can lead to over-optimism and excess investment (Denrell, 2003).[1]

In the following pages, we offer what we believe is the first empirical evidence of Adam Smith's conjecture that young people are prone to selecting occupations with highly skewed incomes, despite the minuscule odds of success. We calculate teenagers' objective probabilities of success in a prototypical winner-take-all market: professional tennis – the only occupation with adequate publicly available premarket and market data.[2] In tennis, one percent of the competitors earn almost half the prize money, which constitutes a more unequal income distribution than in other sports (e.g., see Morales, 2013, Beaton, 2014, or Bednall, 2015) and roughly double the top US income distribution (e.g., Alvaredo et al., 2013). Tennis offers international rankings of potential entrants as young as 13 years of age, which we link to the same players' professional outcomes and earnings later in life. The nature of tennis allows us to isolate individual performance and ability over an entire career, rather than having to decompose individuals' efforts in team environments. We create a longitudinal dataset that tracks *all* male and female players born between 1977 and 1986 from ages 13 to 30 who have been ranked in the U14, U16, U18, or professional rankings. Overall, this produces a sample of 7,242 male and 6,205 female players.

With these data, we estimate each player's predicted lifetime prize money distribution, given his/her global ranking at each age between 13 and 19, as well as the objective probabilities of

[1]For example, by eliminating poorly performing products, mutual funds overstate their performance and understate their risk (Elton et al., 1996).

[2]Golf, the other major lucrative individual-level sport, lacks global junior ranking data.

changing rank from one year to another. The shape of these conditional prize money distributions varies widely across players, with lower-ranked players exhibiting extreme levels of skewness, reflecting the remoteness of the chance of becoming a superstar. We calculate the moments of these distributions and document how they influence teenage players' decisions to continue in the sport or to quit tennis. Our estimations account for player fixed effects, which allows us to control for any unobservable characteristics at the individual level that may otherwise confound our analysis, such as personal or family wealth, preferences, innate ability, and support from local and national tennis federations.

Consistent with economic theory, we find that teenage tennis players are more likely to stay in the sport when they face high mean predicted earnings and a low variance. However, players are also attracted to highly-skewed earnings distributions, much like decision-makers in insurance and financial investment markets and horse race or lottery gamblers. Under a hypothetical scenario of zero earnings skewness, male teenagers would be 23 percent less likely, on average, to continue in the sport. Thus, superstar labor markets appear to entice aspirants of modest abilities with negligible chances of earning positive returns to continue making skill-specific human capital investments. This analysis extends the existing empirical literatures regarding human capital investments and occupational choice (Hartog and Diaz-Serrano, 2014), decision-making regarding skewness, risk and return (O'Donoghue and Somerville, 2018), superstar/winner-take-all markets (Rosen, 1981; Frank and Cook, 2010), and learning about one's abilities (Arcidiacono et al., 2012).

Our dataset also offers a rare opportunity to investigate gender differences in occupational choice in the presence of highly-skewed earnings, since tennis is one of the few sports in which men and women compete in independent labor markets for comparable returns. In contrast to males, female teenagers would be only 5 percent less likely to continue if skewness were to hypothetically fall to zero. A corollary of these findings is that one reason for the lack of female entrepreneurs and chief executives may be women's relatively smaller attraction to skewed earnings distributions, compared to men. These findings remain independent of the expected mean earnings and variance, which have been highlighted as other reasons for women's absence in such positions (e.g., see Jianakoplos and Bernasek, 1998, Croson and Gneezy, 2009, and Berkhout et al., 2010).

Finally, to contextualize our results, we compare the corresponding magnitudes regarding high-stakes career gambles to those from an analogous analysis of low-stakes lotteries, using Rieger et al.'s (2014) data. Low-stakes lottery players exhibit 14-21 times greater preferences for skewness than do tennis players when making career choices. This is as expected since the main motivation for purchasing a $2 lottery ticket is the entertainment value of fantasying about how hitting a jackpot would change one's life (Clotfelter and Cook, 1989). Even so, a significant number of young tennis players – especially males – are similarly seduced by superstars' income levels, but with much greater financial and psychological consequences than buying a lottery ticket.

# 2 Background

## 2.1 Rational Expectations and Winner-Take-All Markets

Two theories have been proposed to explain decision-making under uncertainty in the context of winner-take-all markets. First, Rosen and Sanderson (2001) follow a standard rational expectations utility model to characterize participants in winner-take-all labor markets as learning about their abilities and prospects of success by regularly reassessing the expected value of their lifetime earnings or their chances of employment by, for example, a major-city orchestra or a professional sports team. In our setting, junior tennis players receive weekly tournament performance feedback and year-end rankings relative to their peers as they decide whether to drop out or to continue pursuing a tennis career. In a standard rational expectations expected utility model, an agent weights the sums of the utility values of outcomes multiplied by their respective probabilities (Von Neumann and Morgenstern, 2007). Such a framework is in the spirit of MacDonald's (1988) superstar model in which ability gradually reveals itself over time based on the accumulation of information about one's performances, with superstar earnings providing the proper incentives to enter these professions.

Related job-matching models are provided by Rosen's (1981) superstar model and Miller's (1984) occupational choice model. Rosen and Sanderson (2001) suggest that the "option value of occupational risk-taking" encourages entry, but also limits the risk of social and private losses, akin to the standard value of an option in finance. Stange (2012) estimates "that option value accounts for 14 percent of the total value of the opportunity to attend college for the average high school graduate and is greatest for moderate-aptitude students".

## 2.2 Prospect Theory and Winner-Take-All Markets

Second, whereas expected utility theory assumes that all percentage points of risk are equally important, prospect theory proposes that the values of outcomes of risky prospects be multiplied by decision weights that "measure the impact of events on the desirability of prospects, and not merely the perceived likelihood of these events" (Kahneman and Tversky, 2013, p. 280). As suggested by the Allais paradox, ample evidence indicates that people value very small changes in the probability of big payoffs over larger changes nearer the middle (Savage, 1972; Slovic and Tversky, 1974). Finance scholars have addressed shortcomings of expected utility models in which skewed returns induce risk-seeking behavior with models of loss aversion (Benartzi and Thaler, 1995), salience (Dertwinkel-Kalt and Köster, 2017), and context dependence (Bordalo et al., 2013; O'Donoghue and Somerville, 2018).

Although entrants into superstar markets are often assumed to be risk lovers, the empirical literature on gambling on horse races and lotteries suggests instead that punters are willing to accept a lower expected payoff in return for greater skewness in the payoff distribution. Cook and Clotfelter (1993) and Forrest et al. (2002) find lotto sales to be positively related to the size of the jackpot, but negatively to expected value. Golec and Tamarkin (1998) find horse track bettors to be attracted to the positive skewness of returns offered by low probability, high variance bets, rather than being risk lovers with mean-variance utility functions. Garrett and Sobel (1999, p.88) conclude that "lottery players, like horse race bettors, are risk averse but favor positive skewness".[3]

Our approach is similar to that of Forrest et al. (2002) who find that high headline maximum possible jackpot prizes exert an influence upon lotto demand beyond the effective price. Whereas lotteries entail purely random outcomes, we analyze outcomes in a winner-take-all labor market based on individual abilities. Related empirical analyses of decisions made in the presence of uncertainty about one's own abilities and expected future outcomes investigate whether or not to attend college (Kane, 1994), whether to continue or to dropout once in college (Stange, 2012; Stinebrickner and Stinebrickner, 2012), and the choice of college major (Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2013). Generally, these authors find a reasonable approximation to Bayesian updating based on sequential experiences that provide new information about an individual's match with particular training programs. Note that students learn about their academic interests and abilities semester by semester with the option to drop out and infer their post-college earnings from those academic experiences.

What factors differentiate the circumstances in which decision-makers respond more to the probabilities than to the attractiveness of potential outcomes? There is some evidence that people respond to *positive* feedback close to Bayes' rule, but are less responsive to *negative* feedback (Eil and Rao, 2011). Market efficiency depends upon information, complexity, and human analytical ability. Sobel and Ryan (2008) show that the horse race longshot bias results

[3] For the experimental literature regarding the longshot anomaly and evidence for positive skewness preferences, see Grossman and Eckel (2015).

from casual bettors' reliance on selected and unreliable information, whereas serious bettors and arbitrageurs use both better information and better analyses of it. Camerer and Lovallo (1999) find that optimistic over-entry persists if performance feedback necessary to correct it is noisy, infrequent, and slow, which characterizes the circumstances for excess entry of many entrepreneurs and participants in the creative industries and which distinguishes choices about continuing in college and the college major compared to selecting one's first job. In our context of tennis, players can participate in weekly tournaments to obtain objective performance feedback which is eventually used to update global rankings.

## 2.3 Labor Supply Decisions, Income Risk, and Skewness

A small theoretical and empirical literature analyzes the relationship between income risk, skewness, and career choice (Barth et al., 2017; Hartog and Diaz-Serrano, 2014). One strand of that literature interprets income risk as general labor market uncertainty measured by wage dispersion in different occupations (King, 1974; Johnson, 1977) or via individual wage dispersion over a certain time period (Moore, 1995). The other strand of literature focuses on the decision to invest in education and relates to our paper. Levhari and Weiss (1974) provide a theory for the impact of income risk on educational choice in what has become the standard model for human capital decisions under uncertainty (also see Krebs, 2003). Flyer (1997) concludes that "[s]tudents' uncertainty over their relative abilities across different occupational fields, combined with the variance and skewness of pay distributions in an environment with occupational mobility, indicate that projections of future earnings identified with a profession derive mainly from the right-hand tail of the pay distribution" (also see Harris and Weiss, 1984). Hartog and Diaz-Serrano (2014) summarize the empirical results of both strands of literature and find evidence of a positive compensation for income variance and a negative compensation for skewness. Some of these studies also speak to potential gender differences, concluding that men's earnings fall more than women's due to earnings skewness, i.e., women exhibit less skewness affection than men (see Hartog and Diaz-Serrano, 2014, Tables 7.2 and 7.3).

However, the labor markets studied exhibit much less wage and earnings skewness than our study of a winner-take-all market. For example, Hartog and Vijverberg (2007) find evidence that people are willing to accept lower earnings in return for greater post-schooling earnings skewness, reporting relative earnings skewnesses (i.e., the expected value of the cubed deviation of earnings around its mean, divided by the mean) of merely 0.23 and 0.25 for men and women at the median, respectively, compared to 232 and 81 in our study (also see Berkhout et al., 2010).[4] These vastly higher skewnesses are the defining feature of winner-take-all markets. The existing literature has not estimated how labor market participants behave in such extreme circumstances.

# 3 Theoretical Motivation

To facilitate the analysis of a winner-take-all market, assume that people can choose between a risky career, in which their future earnings are uncertain, and a riskless career, which pays a given amount with certainty. People begin their careers by spending $T$ periods as apprentices, during which they earn zero on either the risky or riskless job. All people initially start on the risky job but can change job any period during the apprenticeship phase. Lifetime earnings on the risky job, $w$, are determined by a person's ordinal ranking according to performance at the end of the apprenticeship phase, which is not known in advance. However, at the end of each apprenticeship period $t$, people learn their current ranking, $r$. Lifetime earnings on the riskless job, $\hat{w}$, are an increasing function of the number of years during the apprenticeship phase spent on the riskless job.

[4] The NBER-CPS data Hartog and Vijverberg (2007) use, like other general labor market data, are top-coded, which biases the estimated skewness downwards. In contrast, our data on tennis players' earnings are not censored.

In each apprenticeship period, person $i$ will choose to continue with the risky career if his/her expected utility in the post-apprenticeship period is greater than that on the riskless career, that is:

$$E_t\big(u(w_i)\big) > u(\hat{w}_i|T-t), \tag{1}$$

where $u$ represents the person's utility function over lifetime income $w_i$. The person's utility function can then be approximated by a Taylor series expansion around the mean of $w$, $\bar{w}$ (see Golec and Tamarkin, 1998):

$$u(w_i) \approx u(\bar{w}) + u'(\bar{w})(w_i - \bar{w}) + \frac{u''(\bar{w})}{2}(w_i - \bar{w})^2 + \frac{u'''(\bar{w})}{6}(w_i - \bar{w})^3. \tag{2}$$

From here, taking expectations at time $t$ and adding an error term produces:

$$E_t\big(u(w_i)\big) = u(\bar{w}) + u'(\bar{w})E_t(w_i - \bar{w}) + \frac{u''(\bar{w})}{2}E_t(w_i - \bar{w})^2 + \frac{u'''(\bar{w})}{6}E_t(w_i - \bar{w})^3 + \epsilon_{it}. \tag{3}$$

Let $\epsilon_{it} \sim U[-e, e]$; then, the probability that the person will continue in the risky career after period $t$ becomes:

$$P(riskycareer_{it}) = P\left(\epsilon_{it} > u(\hat{w}_i) - u(\bar{w}) - u'(\bar{w})E_t(w_i - \bar{w}) - \frac{u''(\bar{w})}{2}E_t(w_i - \bar{w})^2 - \frac{u'''(\bar{w})}{6}E_t(w_i - \bar{w})^3\right)$$

$$= \frac{e - u(\hat{w}_i) + u(\bar{w})}{2e} + \frac{u'(\bar{w})}{2e}E_t(w_i - \bar{w}) + \frac{u''(\bar{w})}{4e}E_t(w_i - \bar{w})^2 + \frac{u'''(\bar{w})}{12e}E_t(w_i - \bar{w})^3. \tag{4}$$

If the person is risk neutral, $u'' = 0$; if the person is skewness neutral, $u''' = 0$. Hence, estimates of a person's squared and cubed deviations of lifetime earnings should be added to an equation for the probability of continuing in the risky career. A significant positive coefficient on the cubed term indicates that the person is skewness-loving.[5]

[5] King (1974), Johnson (1977), and Hartog and Vijverberg (2007) also measure variance and skewness with the standard deviation and the third moment of earnings.

In each period of the apprenticeship phase, a person's expectations depend on their rankings up to that point. Therefore, equation (4) can be rewritten as:

$$P(risky\ career_{it}) = \frac{e - u(\hat{w}_i) + u(\bar{w})}{2e} + \frac{u'(\bar{w})}{2e} E_t(w_i - \bar{w} | \boldsymbol{r}_{it})$$

$$+ \frac{u''(\bar{w})}{4e} E_t\big((w_i - \bar{w})^2 | \boldsymbol{r}_{it}\big) + \frac{u'''(\bar{w})}{12e} E_t\big((w_i - \bar{w})^3 | \boldsymbol{r}_{it}\big), \tag{5}$$

where $\boldsymbol{r_{it}}$ denotes the sequence of person $i$'s rankings up to period $t$.

# 4 Data

## 4.1 Tennis as a Laboratory To Study Winner-Take-All Markets

Our dataset combines four distinct sources for tennis rankings and earnings, all of which are available for both males and females. First, since its inception in 1990, the Tennis Europe Junior Tour publishes year-end rankings for players aged 14 and under (U14) and 16 and under (U16) from around the world who compete in numerous tournaments throughout Europe (for 2018 alone, there are 418 junior tournaments). The results of these tournaments produce the earliest and most comprehensive global rankings for tennis players.[6] The year-end U14 and U16 rankings provide each player's full name, birthday, nationality, and ranking. Second, we access the worldwide rankings for players aged 18 and under (U18), published by the ITF. For an aspiring tennis player, this tour provides the next and final step before entering the professional arena. Similar to Tennis Europe, the ITF rankings include each player's full name, birthday, nationality, and ranking. Third, the complete data regarding players' professional performances

[6] Tennis Europe forms the largest regional federation of the ITF and today manages around 1,200 international tennis events per year (see http://www.tenniseurope.org/page/12173/About-Tennis-Europe). Although European players form the majority on this tour, competitors from 175 countries participate throughout our sample (165 origin countries in the male sample and 156 origin countries in the female sample). Nevertheless, all our results are consistent when focusing on European players only.

come from the respective professional organizations: the ATP for men and the WTA for women. Combining these data sources allows us to construct the first longitudinal data set to analyze career investment decisions by youth within a domain for an entire pool of potential market entrants. (Note that in the youth competitions even those who only play once and lose appear in the respective year-end rankings.)

Earnings data for *all* participants typically are not publicly available for most superstar markets, with the exception of some types of professional athletes (Kahn, 2000). Consequently, existing studies measure success by discrete professional accomplishments – such as winning an Olympic medal, receiving an award, or earning some level of ranking or distinction (e.g., see Brouwers et al., 2012, for tennis and Li et al., 2018, for boxing, taekwondo, and wrestling). Artistic, cultural, and entertainment markets, as well as most team sports, lack objective quality measures and rely on the subjective decisions of coaches, judges, and talent scouts.[7] According to Elferink-Gemser et al. (2011), "to further unravel the mystery of talent, the best way may be to longitudinally follow youth athletes throughout their sport career, from start to adulthood". Rather than retrospectively retracing the steps of superstars, our prospective dataset allows us to study *all* potential market entrants into tennis.

Our analysis focuses on the cohort of players born between 1977 and 1986. This timeframe ensures that we observe players' performances from age 13 (when Tennis Europe rankings became available in 1990) to 30 (since we record players' professional performance until the end of 2016). Our dataset includes every player in the selected cohort who shows up at least once in any of the three sets of rankings, producing a sample of 7,242 male and 6,205 female players. We include every player-year observation in which the player appeared in one of the described rankings in the previous year, leading to 15,810 and 15,078 individual player-year

[7] For example, Caves (2000, p.784) argues that for the creative industries "nobody knows" how consumers will value an art or entertainment product. As a consequence, Krueger (2005), for example, measures rock star quality according to the space devoted to each artist in the Rolling Stone Encyclopedia of Rock and Roll. Judges' scores of music competitions are influenced by arbitrary factors like the order of appearance within the day or week (Ginsburgh and Van Ours, 2003).

observations, respectively. Of those who play professional tennis, the average male and female player in this sample earns \$124,048 and \$99,670 from prize money throughout their career. Out of those who remain active on the professional tours until the age of 30, the average career prize money totals \$2,126,571 (males) and \$1,844,902 (females).

To visualize the winner-take-all nature of tennis, Figure 2 plots annual prize money on the ATP and WTA Tours against players' year-end rankings, illustrating how ranking and earnings are closely related. Because prize money is allocated according to a player's position in each tournament (i.e., the round in which they lose) and declines sharply with tournament prestige, the top players account for the vast majority of total prize money earned. Figure 2 also documents the extremely unequal distribution of earnings in tennis – an artifact that is more pronounced than in any other major sports (also see Beaton, 2014).

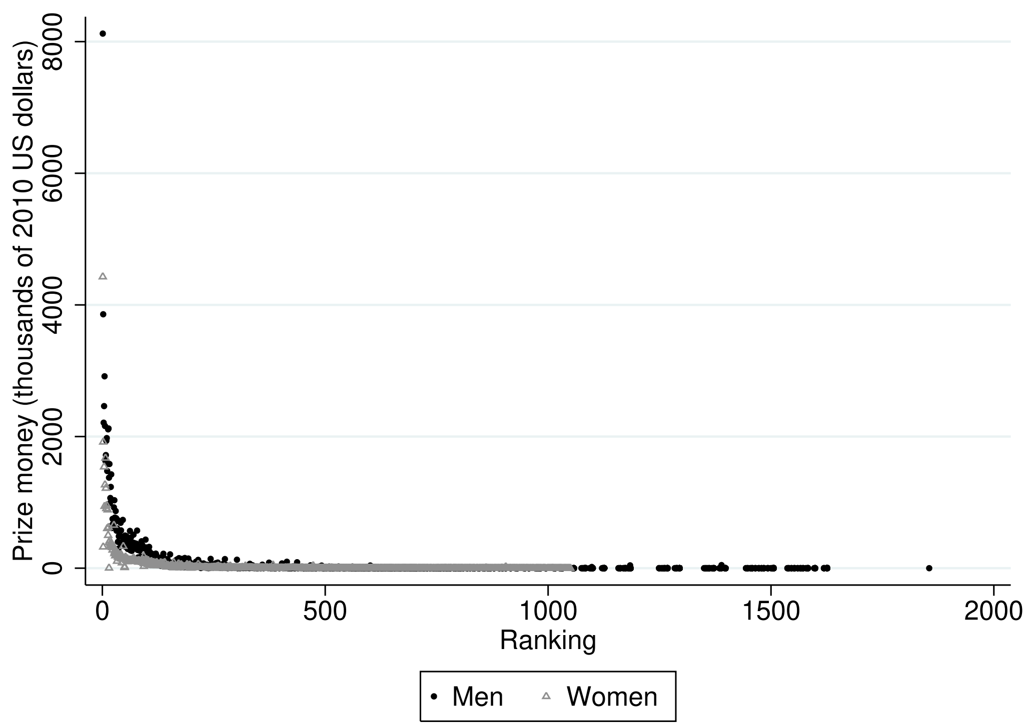


**Figure 2:** Average annual prize money earnings on the ATP (men's) and WTA (women's) Tours.

A young player needs to decide whether to risk entering the professional tennis labor market or to choose a 'safer' career. To do so, they must evaluate and compare the expected distributions of lifetime earnings for the two careers. If players care only about their expected lifetime earnings, the amount of labor supplied should match closely their expected earnings. However, Figure 3 shows that while average lifetime prize money per tournament is highly non-linearly related to junior ranking (in the bottom panel), a player's lifetime number of tournaments is roughly linearly related to ranking (in the top panel). In fact, the ITF's "Player Pathway Review" (2016) concludes that "there are too many players trying to compete on the professional circuit; too few players are breaking even". This anecdotal statement already stands in contrast to Rosen's (1986, p.134) argument, in his review of *The Winner-Take-All Society* (Frank and Cook, 2010), that "few seriously try to enter these [winner-take-all] professions" so that the excess supply "inefficiencies they [Frank and Cook] claim seem to me be greatly exaggerated."

## 4.2 Data Preparation and Descriptive Statistics

Our analysis requires estimates of the lifetime prize money distribution a given player can expect to face at any given age. Details of how we do this are given in the appendix. We first calculate a player's ranking in each year within the cohort of players of the same sex born in the same year, using the relevant U14, U16, U18, and professional rankings. We assume that players observe the probabilities of moving between any two ranks from any year to the next and also that they base their forecasts of future prize money on the distribution of prize money observed in the professional tours in the previous year. Combining the transition probabilities and the observed earnings distributions and considering every possible rank in every future year of a player's career (up to age 30), we then calculate lifetime prize money PDFs for each active player in every year. These values exhibit significant variation. As an example, Figure 4 plots

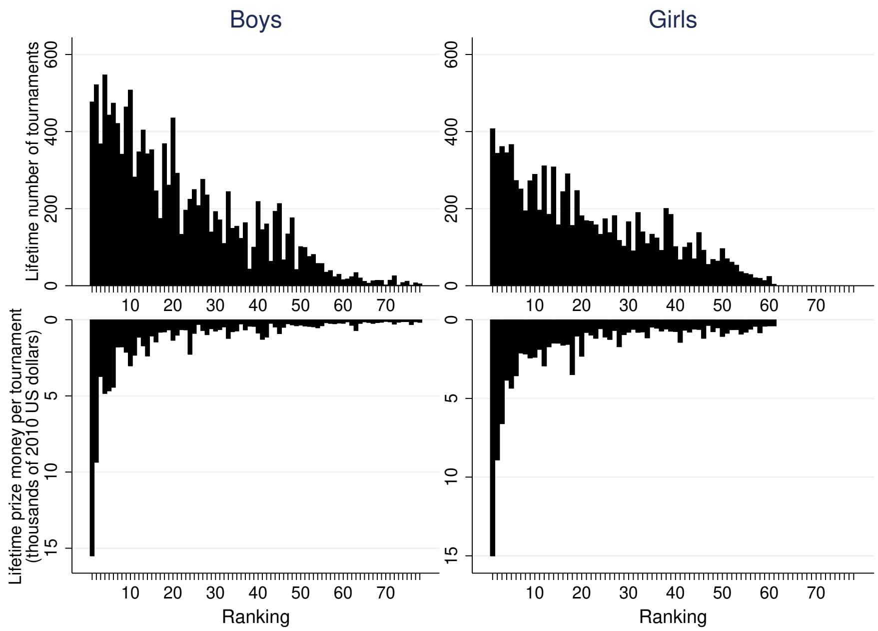


**Figure 3:** Plotting junior rankings (x-axis) against the lifetime number of professional tournaments played (y-axis in top panel) and lifetime prize money earnings per tournament (y-axis in bottom panel).

the distributions faced by 18-year olds in 1997 (the midpoint of our sample) with different rankings. Even those ranked top of their age group face a reasonably high chance of earning very little over their careers. However, those ranked 100 experience a much higher likelihood that they will earn close to zero. This fact is reflected in the skewness coefficients for distributions (calculated as $\frac{E(W_i - \bar{W})^3}{(E(W_i - \bar{W})^2)^{\frac{3}{2}}}$) and reported under each histogram), which are much higher for those ranked 100 than for those ranked one.

Naturally, as players age, their rankings become better predictors of their lifetime earnings. Figures 5 and 6 plot the lifetime prize money distributions for players ranked first among their cohort at each age between 13 and 18. Top-ranked 13-year-olds have a very high chance of making little money over their careers, compared to top-ranked 18-year-olds. Accordingly, their skewness coefficients (listed below each graph) fall with age.

We can calculate three moments of the lifetime earnings distributions for any player $i$ in

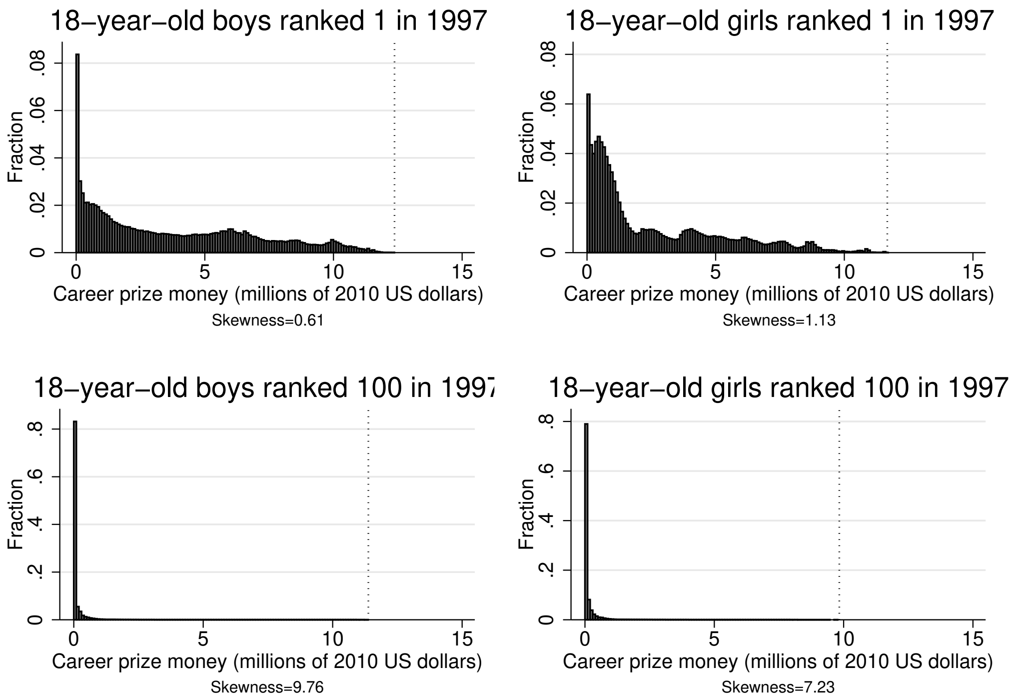


**Figure 4:** Plotting the expected distribution of career prize money for #1's (top panel) and #100's in the U18 rankings.

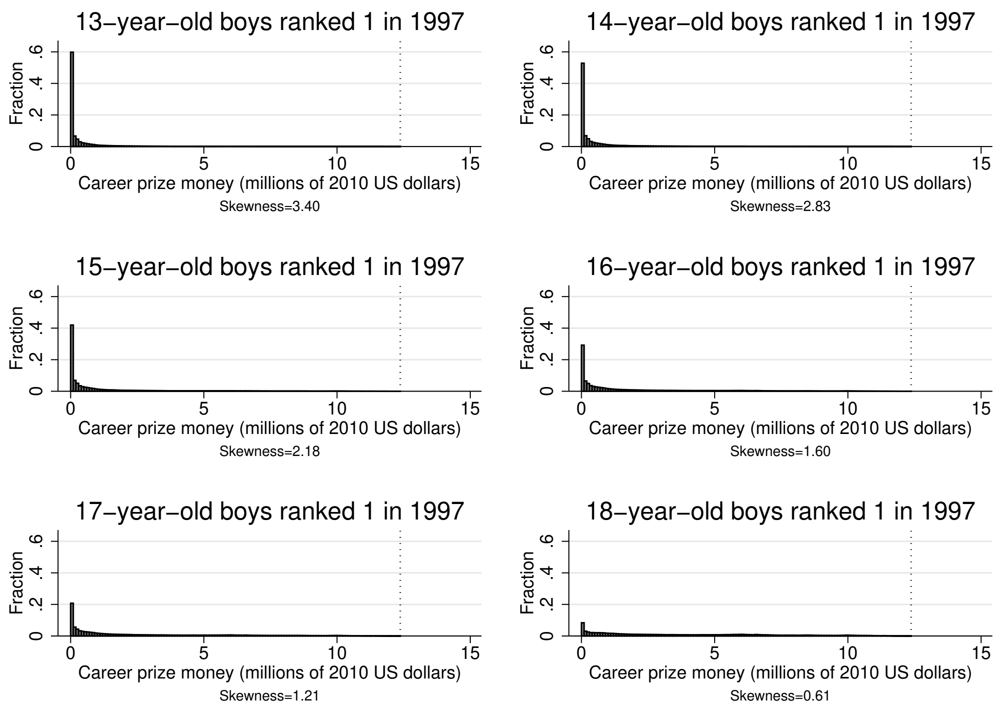


**Figure 5:** Plotting the expected distribution of career prize money for male #1's at the ages 13 until 18.

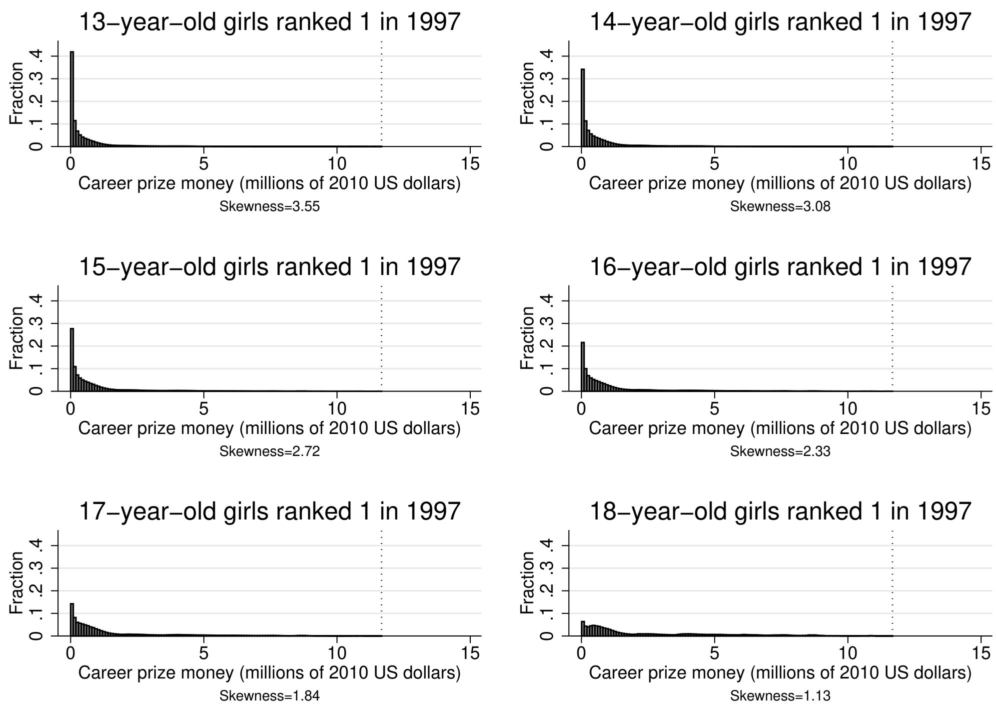


**Figure 6:** Plotting the expected distribution of career prize money for female #1's at the ages 13 until 18.

year $t$, analogous to the three moments from equation (5):

$$E(W_{it}) = \sum_{W} P(W_{it}) W_{it};$$

$$E\big(W_{it} - E(W_{it})\big)^2 = \sum_{W} \big(W_{it} - E(W_{it})\big)^2;$$

$$E\big(W_{it} - E(W_{it})\big)^3 = \sum_{W} \big(W_{it} - E(W_{it})\big)^3. \tag{6}$$

Finally, we need to define what constitutes participation in tennis in our empirical estimations. In our main specifications, we assume a player is active in tennis in a given year if he/she appears in any of the rankings in our dataset (U14, U16, U18, or the professional tours). Table 1 reports the means for the primary variables of interest. Females average slightly higher earnings than males because more males participate in junior tennis than females and they face a higher chance of earning very little over their careers. Male tennis players also face a much greater variance and skewness than female tennis players in their expected earnings. The fact that lagged rankings are worse on average (i.e., take higher values) for males than for females indicates that more males remain in tennis than females, on average.

# 5 Empirical Findings

## 5.1 Empirical Strategy

We regress a binary variable for whether player $i$ was active in year $t$ on the mean, variance, and skewness of his/her expected career prize money, given his/her ranking at the end of year

**Table 1:** Means of key variables for player-year observations in which the player was active in the previous year.

| Variable | Males (1) | Females (2) | (1)-(2) | (p-value for (1)=(2)) |
|---|---|---|---|---|
| Active in tennis | 0.552 | 0.657 | -0.105 | (0.000) |
| Mean of predicted career prize money (millions of 2010 US dollars) | 0.290 | 0.298 | -0.008 | (0.030) |
| Variance of predicted career prize money (trillions of 2010 US dollars squared) | 0.899 | 0.499 | 0.400 | (0.000) |
| Skewness of predicted career prize money (in quintillions of 2010 US dollars cubed) | 5.653 | 2.143 | 3.510 | (0.000) |
| Lagged ranking | 119.150 | 87.029 | 32.121 | (0.000) |
| Age | 16.862 | 16.679 | 0.183 | (0.000) |
| Players | 7,242 | 6,205 | | |
| Number of observations | 15,810 | 15,078 | | |

$t-1$, consistent with equation (5), as follows:

$$P(active_{it}) = \alpha_1 E_t(W_i) + \alpha_2 E_t(W_i - \bar{W})^2 + \alpha_3 E_t(W_i - \bar{W})^3 + \boldsymbol{AGE}_{it}\boldsymbol{\beta} + \gamma_i + \epsilon_{it}, \quad (7)$$

where $\boldsymbol{AGE_{it}}$ represents a full set of age dummies (intended to capture the effects of changes in opportunity cost over a person's teenage years), $\gamma_i$ accounts for player fixed effects, and $\epsilon_{it}$ constitutes a random error term. Throughout our analysis, we analyze males and females separately.

## 5.2 Main Findings for Males

The main results of estimating equation (7) for males are reported in column (1) of Table 2, where we consider all player-year observations of males aged 14-19. For the three moments of prize money, we display the estimated coefficients, the standard errors in parentheses, and the corresponding elasticities in brackets. As predicted by theory, the estimates of $\alpha_1$, $\alpha_2$, and $\alpha_3$ are positive, negative, and positive, respectively. All coefficients are statistically significant on the one percent level. The estimated coefficients imply an elasticity with respect to the mean of prize money of 0.610, an elasticity with respect to the variance of -0.554, and an elasticity with respect to the skewness of 0.228. This means that if the skewness of the prize money distribution were to fall to zero, without a change in the mean and variance, the average male player would be 23 percent less likely to continue playing tennis the following year.

**Table 2:** Results of estimating the participation equation for males.

| | Ages 14-19 | | | Ages 20-28 |
|---|---|---|---|---|
| | (1) | (2) Only Europeans | (3) Costs adjusted | (4) |
| Lagged mean of career prize money (millions of 2010 US dollars) | 1.162*** [0.610] (0.082) | 0.970*** [0.510] (0.109) | 0.904*** [0.179] (0.097) | 0.119*** [0.062] (0.026) |
| Lagged variance of career prize money (trillions of 2010 US dollars squared) | -0.341*** [-0.554] (0.031) | -0.287*** [-0.474] (0.036) | -0.261*** [-0.386] (0.033) | -0.053*** [-0.050] (0.013) |
| Lagged skewness of career prize money (quintillions of 2010 US dollars cubed) | 0.022*** [0.228] (0.002) | 0.020*** [0.205] (0.003) | 0.019*** [0.179] (0.002) | 0.004* [0.014] (0.002) |
| Player- and age-fixed effects[a] | yes | yes | yes | yes |
| $R^2$ | 0.618 | 0.602 | 0.615 | 0.598 |
| Number of observations | 15,810 | 10,581 | 15,810 | 6,009 |

*Notes:* [a] All specifications include a full set of age and player fixed effects. In the second column, prize money is adjusted for the estimated costs of competing. Elasticities at the mean are presented in brackets. Standard errors are presented in parentheses. *, ** and *** denote significance at the 10 percent, 5 percent, and 1 percent level, respectively.

What can explain this skewness finding? Note that, like most of the empirical and experimental studies, we cannot distinguish between preferences for positive skewness and miscalculations about the probability of such outcomes.[8] The latter hypothesis suggests overconfidence: moderately talented players may systematically overestimate their chances of future success, in which case higher skewness in expected earnings could increase the player's desire to continue. In other words, seeing Roger Federer earn more money than before may strengthen a player's desire to continue in professional tennis because he is overestimating the likelihood of being the next Roger Federer. However, the findings from Table 2 suggest that players, parents, and coaches should carefully consider whether to continue investing time and financial resources in tennis and when it is time to quit. Because costly skill-specific human capital investments are indivisible and non-tradeable, the risk associated with longshot gambles for 23 percent of male teenage tennis players may suggest allocative inefficiency.

### 5.3 Main Findings for Females

The results for females are reported in column (1) of Table 3. The coefficients imply an elasticity with respect to the average prize money of 0.276, an elasticity with respect to the variance of -0.146, and an elasticity with respect to the skewness of 0.045. This means that young players would be five percent less likely to stay in tennis, on average, if skewness fell to zero. Note that the corresponding elasticity with respect to skewness for males from column (1) of Table 2 differs in statistical terms from the females' elasticity at the five percent level. In terms of magnitude, the underlying skewness preferences among males are more than five times larger than among females (0.228 versus 0.045).

How do these stark gender differences in skewness preferences compare to the two bodies of research that most relate to our study of entry to winner-take-all markets? One body of research analyzes how educational investment decisions are affected by differences in earnings distri-

[8] For example, Snowberg and Wolfers (2010) cannot distinguish between preference and miscalculation of probabilities explanations for the longshot bias among horse track gamblers.

**Table 3:** Results of estimating the participation equation for females.

| Variable | Ages 14-19 | | | Ages 20-28 |
| --- | --- | --- | --- | --- |
| | (1) | (2) Only Europeans | (3) Costs adjusted | (4) |
| Lagged mean of career prize money (millions of 2010 US dollars) | 0.608*** [0.276] (0.066) | 0.701*** [0.323] (0.077) | 0.237*** [0.017] (0.092) | 0.246*** [0.098] (0.043) |
| Lagged variance of career prize money (trillions of 2010 US dollars squared) | -0.192*** [-0.146] (0.033) | -0.228*** [-0.182] (0.037) | -0.014 [-0.009] (0.048) | -0.104*** [-0.046] (0.023) |
| Lagged skewness of career prize money (quintillions of 2010 US dollars cubed) | 0.014*** [0.045] (0.002) | 0.017*** [0.059] (0.004) | -0.000 [-0.001] (0.006) | 0.001 [0.001] (0.004) |
| Player- and age-fixed effects[a] | yes | yes | yes | yes |
| $R^2$ | 0.643 | 0.633 | 0.641 | 0.558 |
| Number of observations | 15,078 | 9,988 | 15,078 | 4,693 |

*Notes:* [a] All specifications include a full set of age and player fixed effects. In the second column, prize money is adjusted for the estimated costs of competing. Elasticities at the mean are presented in brackets. Standard errors are presented in parentheses. *, ** and *** denote significance at the 10 percent, 5 percent, and 1 percent level, respectively.

butions across occupations. In their review of that literature, Hartog and Diaz-Serrano (2014) report gender differences in skewness affection but of modest magnitudes and with inconsistent findings overall. For example, Hartog and Vijverberg (2007) find men and women to exhibit the same modest skewness preferences based on data with vastly lower skewness levels than the ones we study here. The second literature considers situations in which skewness preferences are found to affect choices regarding lottery-like financial investments, entrepreneurship, and lottery participation. However, we are not aware of separate estimates of men's and women's skewness preferences in the face of payoffs that are as skewed as those faced by tennis players. In a laboratory experiment examining lottery choices, Grossman and Eckel (2015) examine the effect of skewnesses that are more than three times lower than ours (up to 2, using the standard definition, compared to equivalent values in our study of 6.63 for men and 6.08 for women), documenting a 10-15 percent gender gap in skewness preferences. In sum, gender differences in skewness preferences may be even larger in winner-take-all markets than in regular labor markets, especially when skill-specific investments have to be made early in life.

### 5.4 Robustness Checks

Beyond our main results, Tables 2 and 3 also display findings from several alternative specifications. First, since all U14 and U16 tournaments are contested in Europe (and the association is called Tennis Europe itself), it is possible that including non-European players somehow confounds our findings. For example, if only the best non-European players appear in the U14 and U16 rankings, the skewness in lifetime earnings may be overstated. The corresponding results from considering European-born players only are displayed in columns (2) of Tables 2 and 3. Note that the coefficients on the skewness variable remain consistent in terms of statistical relevance and comparable in magnitude: for the males, the respective coefficient on skewness decreases marginally from 0.022 to 0.02, whereas that for females even increases slightly from

0.014 to 0.017.

Second, our main specifications do not account for the costs of competing in professional tennis, which (at least to some degree) vary from player to player. On one hand, since top-ranked players may travel more widely and can afford to contract more support staff than lower-ranked players, our estimates may overstate the skewness in lifetime earnings. On the other hand, if top players have these costs paid for by sponsors or tournament organizers, we may understate the underlying skewness in earnings. (In fact, at the highest-level professional events, hospitality is always included, i.e., hotels are being paid until the respective player loses in the respective week.) To address this, we subtract an estimate of the costs of competing in tennis, provided to aspiring players by the ITF (International Tennis Federation, 2014). This varies by sex, continent, and ranking. Note that, since we have no information on a player's current place of residence, we use their continent of birth as a reference. For the few players who were missing this information, we use the costs faced by those living in Europe, since this is the most common continent of birth. The corresponding results are displayed in column (3) of our main Tables. The skewness elasticity falls only marginally for males (from 0.022 to 0.019), but it loses statistical significance entirely for females. In fact, we derive a precisely estimated null effect. Intuitively, this further highlights gender differences in tennis players' skewness preferences and skewness may not matter at all for female players.

Third and final, columns (4) of Tables 2 and 3 display results from re-estimating equation (7), but using player-year observations for ages 20-28.[9] These players are already on the professional tour and are considering whether to continue or to quit, based on their current ranking. In this case, we derive much smaller skewness effects for both male and female players and, again, the coefficient for women remains statistically indistinguishable from zero. Thus, tennis players are much less influenced by the skewness of income once they have entered the pro-

[9]Note that, although we have rankings and prize money data for each player until age 30, two years of data are lost when constructing the estimated future lifetime prize money distributions.

fessional tennis labor market than when they are aspiring to enter. These results suggest that winner-take-all markets are potentially luring especially young entrants into seeking ‘lottery wins’, at least in the labor market for male tennis players.

## 5.5 Heterogeneity in Skewness Preferences by Stake Size of Gamble

The results discussed above suggest that individuals approach the choice of their career in a manner reminiscent of how they approach lotteries, despite the stakes being considerably higher. Although a person might purchase a lottery ticket for \$2 despite the negative expected return, presumably for the entertainment value or the fantasy of hitting the jackpot, career decisions alter and fundamentally limit a person’s occupational opportunities. How does the magnitude of skewness preferences differ between the choice of career and everyday gambling behavior?

To examine how our estimates relate to those exhibited in smaller lotteries, we compare our results with those obtained from the experimental data collected by Rieger et al. (2014). Rieger et al. (2014) ask participants across college classrooms in 52 countries to give certainty equivalents for a series of hypothetical lotteries in which they stand to gain at most \$10,000 or to lose at most \$100, with an average payoff of around \$800. Although the authors do not examine attitudes towards skewness in particular, estimates of the relationship between the certainty equivalent and skewness can be derived from their data (available at `http://dx.doi.org/10.1287/mnsc.2013.1869`).

In general, person $i$’s risk premium for lottery $j$ can be written as

$$\bar{w}_j - y_{ij} = \frac{u''(\bar{w}_j)}{2u'(\bar{w}_j)} E(w_j - \bar{w}_j)^2 + \frac{u'''(\bar{w}_j)}{6u'(\bar{w}_j)} E(w_j - \bar{w}_j)^3, \tag{8}$$

where $y$ constitutes the person’s certainty equivalent, $w$ represents a possible lottery payoff, and $\bar{w}$ captures the mean payoff. Each of Rieger et al.’s (2014) lotteries features a different mean, variance, and skewness. Therefore, we calculate standardized lotteries by subtracting the mean payout from each lottery and dividing by the standard deviation of the payoffs, so that all payoffs exhibit mean zero and variance one. The risk premium for standardized lottery $j$ can be written as:

$$\frac{-y_{ij}}{\sqrt{E(w_j)^2}} = \frac{u''(0)}{2u'(0)} + \frac{u'''(0)}{6u'(0)} E\left(\frac{w_j}{\sqrt{E(w_j)^2}}\right)^3. \tag{9}$$

The final term in equation (9) represents the skewness of the standardized lottery. Equation (9) implies that the standardized risk premia are linearly related to skewness. Therefore, we estimate this relationship using OLS, allowing for person-fixed effects. The slope coefficients provide estimates of a person's willingness to pay for a unit of skewness, or skewness preferences (which will be positive if people are skewness-loving) and the intercept provides an estimate of a person's risk preferences – specifically the Arrow-Pratt measure of risk aversion divided by two (which will be positive if people are risk averse).

Comparing equation (5) with equation (9), it is clear that the ratio of the coefficient on the skewness and the coefficient on the mean in the former are comparable with the coefficient on skewness in the latter. Both provide an estimate of a person's preferences for skewness, *relative* to their preferences for the mean (i.e., the third derivative of their utility function, divided by six times the first derivative). However, since the career "lottery" in our tennis study is over such a vastly greater amount than in the experimental studies (an average lifetime payoff of around \$300,000, compared to \$800 in Rieger et al., 2014), the utility functions have been linearized around different points in the domain and the two sets of estimates provide an indication of how important skewness is to people when assessing gambles of different magnitudes.

Table 4 presents the slope coefficients when equation (9) is estimated using Rieger et al.'s (2014) data (in column 2), alongside the comparable parameters from the tennis data (in column 1). The Rieger et al. (2014) data reveal a relative skewness preference for men of 0.404 and a relative skewness preference for women of 0.473, which are 21 and 14 times larger than the corresponding estimates from the tennis career decision, respectively. Note that, although we found in Tables 2 and 3 that the overall skewness elasticity was smaller for females than for males, since female players were also much less sensitive to the mean, their *relative* skewness preferences are larger than males', as also found in the Rieger et al.'s (2014) data. Hence, skewness preferences carry a much larger influence on the decision to purchase a lottery ticket than on the decision to enter a risky career. This indicates that people focus on the mean and variance of the potential outcomes when a decision is life-changing and are relatively less influenced by long-shot outcomes – even though a preference for low-probability, high-return outcomes still exerts a sizeable effect on behavior, as seen in Tables 2 and 3.

In Table 4, we also combine observations for males and females within each dataset and allow interactions with age and the GDP quintile of a person's country of birth. With these estimations, we explore whether preferences for skewness differ according to the level of development of a player's country of origin. The latter serves as a proxy for a person's own level of income, which is unavailable in both the tennis and Rieger et al.'s (2014) data. We find more variation in skewness preferences in our data than in Rieger et al.'s (2014) data, consistent with the idea that differences across groups only manifest themselves when the stakes are high. Relative skewness preferences are stable across ages in the low-stakes setting, but dip significantly at age 17 – the point at which many teenagers have to choose whether to go to college – in the tennis data. Although relative skewness preferences are lowest in poor countries when modest-sized lotteries are considered, we find a clear U-shaped relationship in the career choice setting. Young players from the poorest and richest countries appear relatively more attracted to highly-skewed earnings than those from middle-income nations. Nevertheless, these results should be interpreted carefully since a player's country of origin can of course only serve as an imprecise measure of a player's opportunity cost for remaining in tennis. For example, it is well-known that many players, especially from poorer countries (e.g., the post-Soviet countries), move to other countries (e.g., the US, Spain, or other traditional Western European tennis nations) for training opportunities and funding.

**Table 4:** Relative skewness preferences by demographic group.

| Population group | Decision | |
|---|---|---|
| | Continue tennis | Low-stakes lotteries (Rieger et al., 2014) |
| | (1) | (2) |
| Men | 0.019*** | 0.404*** |
| | (0.001) | (0.008) |
| Women | 0.033*** | 0.473*** |
| | (0.005) | (0.007) |
| Age 14 | -0.021 | – |
| | (0.075) | |
| Age 15 | 0.022*** | – |
| | (0.005) | |
| Age 16 | 0.055 | 0.465*** |
| | (0.036) | (0.124) |
| Age 17 | -0.011 | 0.468*** |
| | (0.030) | (0.033) |
| Age 18 | 0.027*** | 0.441*** |
| | (0.009) | (0.010) |
| Age 19 | 0.024*** | 0.440*** |
| | (0.002) | (0.006) |
| Birth country GDP/capita quintile 1 | 0.031*** | – |
| | (0.009) | |
| Birth country GDP/capita quintile 2 | 0.029*** | 0.273*** |
| | (0.006) | (0.040) |
| Birth country GDP/capita quintile 3 | 0.013* | 0.401*** |
| | (0.007) | (0.013) |
| Birth country GDP/capita quintile 4 | 0.018*** | 0.472*** |
| | (0.003) | (0.009) |
| Birth country GDP/capita quintile 5 | 0.025*** | 0.438*** |
| | (0.002) | (0.008) |

*Notes:* Values in the first column constitute ratios of the coefficients on the skewness of predicted prize money and the mean of predicted prize money from a regression of equation (9), using ages 14-19. Values in the second column represent coefficients on the skewness of a lottery from a regression of equation (9). GDP per capita values are measured in 2005 and taken from the World Bank Group (2012). Standard errors are displayed in parentheses. Rieger et al. (2014) do not include anyone aged under 16 in their study and do not have anyone aged under 20 from a GDP/capita quintile 1 country. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent level, respectively.

# 6 Conclusion

*The Winner-Take-All Society* (2010), Frank and Cook's best-selling book, turned Adam Smith's "career lotto" conjecture into a national conversation. But even before that, economists critiqued both the lack of a theoretical explanation for potential inefficiencies in what has become known as winner-take-all markets, as well as the absence of comprehensive data with which to evaluate the extent of the inefficiencies (Rosen, 1986; Galbraith, 1995). Indeed, potential entrants to most winner-take-all markets see only the glorious lives of the winners, such as Jeff Bezos, J.K. Rowling, Beyoncé, or LeBron James, and not the masses who tried and failed.

In this paper, we aim to provide two contributions. First, we provide a simple theoretical motivation for the "career lotto" hypothesis, i.e., labor markets with strongly skewed earnings distributions may attract more entrants than labor markets that are otherwise comparable in expected earnings' mean and variance. Second, we overcome sample selection and survivorship bias problems by assembling a unique data set of the *entire* pool of potential market entrants into professional tennis, a typical winner-take-all market. The individual-level performance measures of tennis (i.e., global rankings) allow us to establish objective pre-market ability measures and combine them with the same players' lifetime earnings in tennis. To our knowledge, this constitutes the first such longitudinal data set available to estimate decision-making by all potential labor market entrants.

We find that skewness in the prize money distributions accounts for 23 percent of the male teenage players who continue to pursue professional tennis careers. These findings are consistent with the hypothesis that the presence of superstars encourages modestly-talented people to make longshot career gambles. Quite strikingly, merely five percent of the teenage female players exhibit such skewness preferences. These results may (at least in part) be able to explain other phenomena, such as the lack of women in the upper echelons of businesses and among entrepreneurs or other prominent manifestations of winner-take-all markets.

In sum, this paper provides some of the first empirical evidence of whether and how labor supply in a winner-take-all market can systematically be driven by the earnings of the very best – independent of the mean and variance of a person's expected earnings. From a societal perspective, this may mean an oversupply of labor, although such a value judgement would require further knowledge about the extent to which individuals exhibit cognitive errors, such as overconfidence or time inconsistency in their preferences. We hope future research can build on our findings to better understand how winner-take-all markets function.

# References

Alvaredo, F., Atkinson, A. B., Piketty, T., and Saez, E. (2013). The top 1 percent in international and historical perspective. *The Journal of Economic Perspectives*, 27(3):3–20.

Arcidiacono, P., Hotz, V. J., and Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16.

Arrow, K. J. (1965). Aspects of the theory of risk-bearing (yrjo jahnsson lectures). *Yrjo Jahnssonin Saatio, Helsinki*.

Baker, J., Cobley, S., Schorer, J., and Wattie, N. (2017). *Routledge handbook of talent identification and development in sport*. Taylor & Francis.

Barberis, N. and Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, 98(5):2066–2100.

Barseghyan, L., Molinari, F., O'Donoghue, T., and Teitelbaum, J. C. (2013). The nature of risk preferences: Evidence from insurance choices. *American Economic Review*, 103(6):2499–2529.

Barth, D., Shore, S. H., and Jensen, S. T. (2017). Identifying idiosyncratic career taste and skill with income risk. *Quantitative Economics*, 8(2):553–587.

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

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

Benartzi, S. and Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, 110(1):73–92.

Berkhout, P., Hartog, J., and Webbink, D. (2010). Compensation for earnings risk under worker heterogeneity. *Southern Economic Journal*, 76(3):762–790.

Berri, D. (2005). *Sports Economics*. Worth Publishers.

Bolton, P., Santos, T., and Scheinkman, J. A. (2016). Cream-skimming in financial markets. *The Journal of Finance*, 71(2):709–736.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and asset prices. *American Economic Review*, 103(3):623–28.

Boyer, B., Mitton, T., and Vorkink, K. (2009). Expected idiosyncratic skewness. *The Review of Financial Studies*, 23(1):169–202.

Brouwers, J., De Bosscher, V., and Sotiriadou, P. (2012). An examination of the importance of performances in youth and junior competition as an indicator of later success in tennis. *Sport Management Review*, 15(4):461–475.

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4):1092–1118.

Brynjolfsson, E. and McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1):306–318.

Caves, R. E. (2000). *Creative industries: Contracts between art and commerce*. Number 20. Harvard University Press.

Chen, J. S., Croson, D. C., Elfenbein, D. W., and Posen, H. E. (2018). The impact of learning and overconfidence on entrepreneurial entry and exit. *Forthcoming in Organization Science*.

Clotfelter, C. T. and Cook, P. J. (1989). The demand for lottery products. Technical report, National Bureau of Economic Research, Inc.

Collier, B. L., Schwartz, D., Kunreuther, H. C., and Michel-Kerjan, E. O. (2017). Risk preferences in small and large stakes: Evidence from insurance contract decisions. Technical report, National Bureau of Economic Research.

Cook, P. J. and Clotfelter, C. T. (1993). The peculiar scale economies of lotto. *The American Economic Review*, 83:634–643.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474.

Denrell, J. (2003). Vicarious learning, undersampling of failure, and the myths of management. *Organization Science*, 14(3):227–243.

Dertwinkel-Kalt, M. and Köster, M. (2017). Salient compromises in the newsvendor game. *Journal of Economic Behavior & Organization*, 141:301–315.

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–38.

Elferink-Gemser, M. T., Jordet, G., Coelho-E-Silva, M. J., and Visscher, C. (2011). The marvels of elite sports: How to get there? *British Journal of Sports Medicine*, 45(9):683.

Elton, E. J., Gruber, M. J., and Blake, C. R. (1996). Survivor bias and mutual fund performance. *The Review of Financial Studies*, 9(4):1097–1120.

Ericsson, K. A., Hoffman, R. R., Kozbelt, A., and Williams, A. M. (2018). *The Cambridge handbook of expertise and expert performance*. Cambridge University Press.

Flyer, F. A. (1997). The influence of higher moments of earnings distributions on career decisions. *Journal of Labor Economics*, 15(4):689–713.

Forrest, D., Simmons, R., and Chesters, N. (2002). Buying a dream: Alternative models of demand for lotto. *Economic Inquiry*, 40(3):485–496.

Frank, R. H. and Cook, P. J. (2010). *The winner-take-all society: Why the few at the top get so much more than the rest of us*. Random House.

Galbraith, J. K. (1995). The winner takes all... sometimes. *Harvard Business Review*, 73(6):44–45.

Garrett, T. A. and Sobel, R. S. (1999). Gamblers favor skewness, not risk: Further evidence from united states' lottery games. *Economics Letters*, 63(1):85–90.

Gibson, J., Johnson, P., and Dimita, G. (2015). The business of being an author: A survey of authors earnings and contracts. *Queen Mary University of London, https://www. alcs. co. uk/Documents/Final-Report-For-Web-Publication-(2). aspx*.

Ginsburgh, V. A. and Van Ours, J. C. (2003). Expert opinion and compensation: Evidence from a musical competition. *American Economic Review*, 93(1):289–296.

Golec, J. and Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *Journal of Political Economy*, 106(1):205–225.

Gollier, C. (2001). Wealth inequality and asset pricing. *Review of Economic Studies*, 68(1):181–203.

Green, T. C. and Hwang, B.-H. (2012). Initial public offerings as lotteries: Skewness preference and first-day returns. *Management Science*, 58(2):432–444.

Grossman, P. J. and Eckel, C. C. (2015). Loving the long shot: Risk taking with skewed lotteries. *Journal of Risk and Uncertainty*, 51(3):195–217.

Guglielmo, C. (2014). Mobile apps won't lead to riches for most developers. *Forbes*, January 13, 2014.

Harris, M. and Weiss, Y. (1984). Job matching with finite horizon and risk aversion. *Journal of Political Economy*, 92(4):758–779.

Hartog, J. and Diaz-Serrano, L. (2014). Schooling as a risky investment: A survey of theory and evidence. *Foundations and Trends® in Microeconomics*, 9(3–4):159–331.

Hartog, J. and Vijverberg, W. P. (2007). On compensation for risk aversion and skewness affection in wages. *Labour Economics*, 14(6):938–956.

International Tennis Federation (2014). ITF pro circuit review stage one: Data analysis. `http://www.itftennis.com/media/194256/194256.pdf`.

Jehiel, P. (2018). Investment strategy and selection bias: An equilibrium perspective on overoptimism. *American Economic Review*, 108(6):1582–97.

Jianakoplos, N. A. and Bernasek, A. (1998). Are women more risk averse? *Economic inquiry*, 36(4):620–630.

Johnson, W. R. (1977). Uncertainty and the distribution of earnings. In *The distribution of economic well-being*, pages 379–396. NBER.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *Journal of Economic Perspectives*, 14(3):75–94.

Kahneman, D. and Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pages 99–127. World Scientific.

Kane, T. J. (1994). College entry by blacks since 1970: The role of college costs, family background, and the returns to education. *Journal of Political Economy*, 102(5):878–911.

Kimball, M. S. (1990). Precautionary saving in the small and in the large. *Econometrica*, 58(1):53–73.

King, A. G. (1974). Occupational choice, risk aversion, and wealth. *Industrial and Labor Relations Review*, 27(4):586–596.

Krebs, T. (2003). Human capital risk and economic growth. *The Quarterly Journal of Economics*, 118(2):709–744.

Krueger, A. B. (2005). The economics of real superstars: The market for rock concerts in the material world. *Journal of Labor Economics*, 23(1):1–30.

Levhari, D. and Weiss, Y. (1974). The effect of risk on the investment in human capital. *American Economic Review*, 64(6):950–63.

Li, P., De Bosscher, V., Pion, J., Weissensteiner, J. R., and Vertonghen, J. (2018). Is international junior success a reliable predictor for international senior success in elite combat sports? *European journal of sport science*, 18(4):550–559.

Lubinski, D., Benbow, C. P., and Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science*, 25(12):2217–2232.

MacDonald, G. (1988). The economics of rising stars. *American Economic Review*, 78(1):155–66.

Miller, R. A. (1984). Job matching and occupational choice. *Journal of Political Economy*, 92(6):1086–1120.

Moore, M. J. (1995). Unions, employment risks, and market provision of employment risk differentials. *Journal of Risk and Uncertainty*, 10(1):57–70.

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

O'Donoghue, T. and Somerville, J. (2018). Modeling risk aversion in economics. *Journal of Economic Perspectives*, 32(2):91–114.

Philippon, T. and Reshef, A. (2012). Wages and human capital in the US finance industry: 1909–2006. *The Quarterly Journal of Economics*, 127(4):1551–1609.

Rieger, M. O., Wang, M., and Hens, T. (2014). Risk preferences around the world. *Management Science*, 61(3):637–648.

Rosen, S. (1981). The economics of superstars. *The American Economic Review*, 71(5):845–858.

Rosen, S. (1986). The theory of equalizing differences. *Handbook of Labor Economics*, 1:641–692.

Rosen, S. and Sanderson, A. (2001). Labour markets in professional sports. *The Economic Journal*, 111(469):47–68.

Savage, L. J. (1972). *The Foundations of Statistics*. Courier Corporation.

Slovic, P. and Tversky, A. (1974). Who accepts savage's axiom? *Systems Research and Behavioral Science*, 19(6):368–373.

Snowberg, E. and Wolfers, J. (2010). Explaining the favorite-long shot bias: Is it risk-love or misperceptions? *Journal of Political Economy*, 118(4):723–746.

Sobel, R. S. and Ryan, M. E. (2008). Unifying the favorite-longshot bias with other market anomalies. *Handbook of Sports and Lottery Markets, ed. D Hausch, W Ziemba*, pages 137–60.

Stange, K. M. (2012). An empirical investigation of the option value of college enrollment. *American Economic Journal: Applied Economics*, 4(1):49–84.

Stinebrickner, R. and Stinebrickner, T. R. (2013). A major in science? Initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies*, 81(1):426–472.

Stinebrickner, T. and Stinebrickner, R. (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, 30(4):707–748.

Von Neumann, J. and Morgenstern, O. (2007). *Theory of games and economic behavior*. Princeton University Press.

World Bank Group (2012). *World Development Indicators 2012*. World Bank Publications.

# A Appendix: Construction of Lifetime Prize Money Distributions

In each year, we calculate a player's ranking within the cohort of players of the same sex born in the same year, $b$, using the relevant U14, U16, U18, and professional rankings. For players who have a ranking in more than one age category in a year (e.g., a 16 year-old who plays in both the U16 and U18 categories), the best age-specific ranking is taken. This allows us to derive a cohort-specific ranking for each player every year over their entire career.[10]

We assume that a player's total future prize money, $W$, is determined solely by their age-specific ranking, $r$, in each future year: an assumption that seems reasonable given the tight relationship in visualized in Figure 3. In that case:

$$W_{it} = W(r_i) = \sum_{v=1}^{b+30} w\big(r_{i(t+v)}\big). \tag{10}$$

To simplify our analysis, our empirical estimations assume that players only receive prize money between the ages of 19 and 30. Since players do not know the future prize money distribution, we assume that they use the distribution of prize money on the professional tour (i.e., WTA and ATP) in the previous year as a reference point. Therefore, prize money $v$ years in the future is given by:

$$w\big(r_{i(t+v)}\big) = w\big(r_{t+v}, a_{it} + v, s_i, t-1\big), \tag{11}$$

where the right-hand side gives the prize money of a player ranked $r$ among a cohort of players aged $a$ and of sex $s$ in year $t-1$. Equation 7 says that, for example, a 13-year old boy in 1992 who expects to be ranked 3$^{rd}$ among his birth cohort in 10 years' time would expect to earn the same amount (adjusted for inflation) as the 3$^{rd}$ best 23 year-old man in the professional tour in 1991. To simplify our computations, prize money is rounded to the nearest \$1,000.

The PDF for the lifetime prize money distribution gives the probability of each lifetime earnings amount arising, which is equal to the probability of a given sequence of rankings over a player's lifetime:

$$P\big(W_{it}\big) = P\big(r_{i(t+30-a)}, r_{i(t+29-a)}, \ldots, r_{i(t+1)}\big). \tag{12}$$

Players cannot possibly know the probability of a given sequence of rankings, since there are countless such sequences. However, it seems reasonable to assume that players would know how the probability of any given ranking in the following year is related to their current age-

[10]In some instances, a player may only be listed in a higher category. For example, a 16-year old may only choose to compete in U18 tournaments in a given year and would therefore only appear in the U18 rankings, not the U16 rankings. In these cases (which account for 19 percent of all player-year observations), we use a person's ranking among players of the same age in the higher category only. Thus, there could be two number one 16-year olds in a given year: one who competed in the U16 and another who competed only in the U18. Nevertheless, our results are virtually unchanged when we exclude the observations where players competed only in a higher category.

specific ranking. If we assume that the probability of any ranking is determined solely by a player's age-specific ranking in the previous year, we can simplify the previous expression as follows:

$$P\big(W_{it}\big) = P\big(r_{t+v}|r_{t+v-1}, a_{it} + v - 1, s_i\big)P\big(r_{t+v-1}|r_{t+v-2}, a_{it} + v - 2, s_i\big)...P\big(r_{t+1}|r_{it}, a_{it}, s_i\big). \tag{13}$$

The year-to-year transition probabilities are calculated by comparing the rankings of all players of a given age and sex in one year and the next. For example, we assume the 13 year-old boy referred to above knows – and bases his decisions on – the probabilities of a boy moving between any two ranking places between ages 13 and 14. To simplify the computations, we group players into 10-rank bands between rankings 50 and 500 and 100-ranking bands above ranking 500.