



Citation for published version:

Grove, W, Jetter, M & Papps, KL 2021, 'Career lotto? Labor supply in a superstar market', *Journal of Economic Behavior and Organization*, vol. 183, pp. 362-376. <https://doi.org/10.1016/j.jebo.2020.12.026>

DOI:

[10.1016/j.jebo.2020.12.026](https://doi.org/10.1016/j.jebo.2020.12.026)

Publication date:

2021

Document Version

Peer reviewed version

[Link to publication](#)

Publisher Rights

CC BY-NC-ND

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Career lotto? Labor supply in a superstar market*

Wayne Grove[†]

Michael Jetter[‡]

Kerry L. Papps[§]

November 9, 2020

Abstract

Are people prone to selecting occupations with highly skewed income distributions, irrespective of the mean and variance of their expected earnings? This paper introduces a basic theoretical framework to understand labor supply decisions in winner-take-all markets. Assembling a comprehensive longitudinal dataset of potential tennis professionals (a typical superstar market), we use objective and publicly known rankings to construct earnings projections *before* each player decides whether to enter that labor market. Player- and cohort-age fixed effects account for unobservable factors at the individual level and within a given cohort and year. We find prospective tennis professionals are attracted to highly skewed earnings distributions, independent of mean and variance. Hypothetically, if skewness in prize money fell to the level of the overall US labor market, males would be 7.2 percent and females 2.6 percent less likely to continue pursuing a professional tennis career, on average. Superstar labor markets may therefore systematically encourage those with modest talents to pursue long-shot careers.

JEL Classifications: D81; J22; J24; J31; J44; L83

Keywords: superstar markets; winner-take-all markets; labor supply; human capital; skewness preferences

*We are grateful to Lowell Taylor, 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, the Gijón Conference on Sports Economics, and the Royal Economic Society annual meetings, as well as attendees at the research seminars at Curtin University, Le Moyne 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.

[†]Le Moyne College, 1419 Salt Springs Road, Syracuse 13214, NY, USA. Email: grovewa@lemoyne.edu.

[‡]University of Western Australia, IZA (Bonn), and CESifo (Munich); 35 Stirling Highway, Crawley 6009, Australia. Email: mjetter7@gmail.com

[§]University of Bath, Bath, BA2 7AY, United Kingdom. Email: K.L.Papps@bath.ac.uk

“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. Adam Smith’s conjecture is that the astonishing fame and fortune of superstars – like today’s Beyoncé, Jeff Bezos, LeBron James, and J.K. Rowling – encourages *too many* entrants into the occupations they dominate, causing a socially inefficient allocation of resources (also see [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. The consequences for policymakers are significant: whereas the latter argument is consistent with a well-functioning labor market, the former argument implies a market failure that may justify government intervention and a re-thinking of income structures in superstar labor markets.

A highly skewed income distribution constitutes the defining feature of superstar or winner-take-all markets. The ability to compete in such markets, though, only reveals itself by on-the-job talent discovery, such as the success of a computer app, 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, which for the domain we examine are estimated to cost from US\$150,000-US\$400,000 from age 5-18.¹ Consequently, most potential entrants face negative rates of return,

¹To prepare a junior player for the professional tennis tour, Patrick McEnroe, the general manager of the

with mean earnings being dwarfed by both the expenses associated with pre-market development and market participation, as well as the opportunity costs incurred by comparatively low earnings in ancillary careers (e.g., see [Barberis and Huang, 2008](#)).

Then how can we explain people’s propensity to enter superstar labor 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 – akin to a “jackpot effect” – in horse race and lottery bets ([Cook and Clotfelter, 1993](#); [Garrett and Sobel, 1999](#); [Forrest et al., 2002](#)), financial investments ([O’Donoghue and Somerville, 2018](#)), entrepreneurship decisions ([Chen et al., 2018](#)), and insurance choices ([Barseghyan et al., 2013](#)). Ample evidence from behavioral economics suggests individuals make decisions under uncertainty based on the attractiveness of outcomes, not just their underlying probabilities ([Kahneman and Tversky, 2013](#); [Dertwinkel-Kalt and Köster, 2020](#)).

In the context of labor markets, however, studies on highly skewed earnings distributions have remained elusive.² The most basic dilemma is that “data are not available to calculate meaningful success probabilities for potential entrants” ([Rosen and Sanderson, 2001](#)). In his economics of rock & roll article, [Krueger \(2005\)](#) writes “[t]he superstar model has proved difficult to test empirically because an objective measure of star quality for popular musicians is hard to define and even harder to quantify.”³ The same applies to most other superstar domains, except in sports where individual productivity, to varying degrees, can be measured objectively.

US Tennis Association, estimated US\$150,000, and the British Lawn Tennis Association about £250,000 (or US\$400,000) from age 5-18 (see [Rossing, 2011, 2016](#), and [Gloster, 2015](#)).

²Existing works usually analyze professions that exhibit much smaller earnings skewnesses. For example, [Hartog and Vijverberg \(2007\)](#) find 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](#)). Also note that the data [Hartog and Vijverberg \(2007\)](#) use, like other general labor market data, are top-coded, which biases the estimated skewness downwards.

³[Krueger \(2005\)](#) does so by counting the millimeters of print columns (including photos) for each artist in the The Rolling Stone Encyclopedia of Rock & Roll.

Furthermore, career earnings are rarely available for *all* superstar market participants. However, calculating odds of success requires objective measures of ability and earnings potential for a *comprehensive* database of potential market entrants – rather than an *a posteriori* selected subsample of the successful few (Denrell, 2003; Jehiel, 2018). We have identified such a domain – tennis – and assembled the first dataset that meets these criteria.

In the following pages, we first introduce a simple theoretical model to illustrate the decision of a potential entrant into a winner-take-all labor market. Our focus lies on the first three moments of a player’s expected earnings, conditional on their current ranking: mean, variance, and skewness of what s/he may be able to expect on the professional stage in the winner-take-all labor market, compared with a ‘riskless’ alternative career. We then offer what we believe is the first comprehensive, longitudinal analysis of labor supply in a winner-take-all market by calculating teenagers’ objective probabilities of success in professional tennis. With one percent of the competitors earning almost half the prize money, tennis features a more unequal income distribution than any other major sport (e.g., see Morales, 2013, Beaton, 2014, or Bednall, 2015) and roughly double that of the overall US income distribution (e.g., Alvaredo et al., 2013). Regardless of their innate talent, prospective professional players are required to devote an enormous amount of time to practice throughout their childhoods and teenage years, at the expense of other educational activities. Tennis offers transparent, 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 as an individual sport allows us to cleanly isolate performance, rather than having to decompose and conjecture an individual’s contribution from team performances. Our longitudinal pipeline dataset tracks *all* male and female players born between 1977 and 1986 from ages 13 to 30 who have participated in at least one event in the U14, U16, U18, or professional tours. Overall, this produces a universe of 7,242 male and 6,205 female players.

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 changing rank from one year to another.⁴ The shape of these conditional prize money distributions varies widely across players, with lower-ranked players exhibiting more extreme levels of skewness, reflecting the remoteness of the chance to become a superstar. We calculate the moments of these distributions and explore whether and how they are able to explain a teenage player's decision to continue with or quit tennis. Our longitudinal dataset allows us to account for player fixed effects, thereby controlling for unobservable characteristics at the individual level that may affect participation. Thus, characteristics that do not change (or only change little) between the ages of 13 and 19 for a particular individual are held constant, such as individual backgrounds (financial and otherwise) and preferences, innate ability, opportunity costs, enjoyment from and personal preferences related to playing tennis, as well as support from local and national tennis federations. Our analysis also accounts for cohort-age fixed effects, removing any statistical variation that is unique to a particular cohort and year, such as a general increase in participation due to the emergence of a new superstar. The variation that remains comes from year-to-year changes in a player's junior ranking, which – we argue – solely reflects new information about the person's relative position among all prospective professional players, as the player and his/her competitors mature physically and their likely talent as adults is gradually revealed. Hence, we test whether players respond to the acquisition of new information in precisely the way [Rosen and Sanderson \(2001\)](#) speculated. We examine whether a given player is more or less likely to quit tennis in the year after discovering, for example, that he/she is ranked lower than previously thought and therefore faces lower mean lifetime earnings but also a more skewed earnings distribution.

Consistent with economic theory, we find teenage players are more likely to continue when

⁴We realize that prize money is not the only reward for success in professional tennis as endorsements, appearance fees, and other amenities are often given out to players. Unfortunately, it remains impossible to collect this information for *every* player, largely because these data remain unreported. Nevertheless, if anything, these amenities are available to a substantially larger degree for the top players, i.e., the inequality in professional tennis may be larger than that derived from prize money alone. Thus, one could see our results as lower bound estimates.

they face high mean predicted earnings and a low variance. However, players are also attracted to highly skewed earnings distributions, much like horse race or lottery gamblers. Under a hypothetical scenario in which skewness fell to the level of the overall US labor market, male and female teenagers would be 7.2 percent and 2.6 percent less likely, on average, to continue in tennis. Thus, superstar labor markets may systematically entice aspirants of modest abilities, who face negligible chances of earning positive returns, to continue making skill-specific human capital investments.

Our paper aims to contribute to the literature in three ways. First, our theoretical model and empirical findings provide an understanding of labor supply in winner-take-all markets, complementing [Rosen's \(1981\)](#) superstar model and [MacDonald's \(1988\)](#) rising star model. Substantial empirical research has analyzed superstar labor markets, for example among professional athletes (see [Berri, 2005](#)), musicians ([Krueger, 2005](#)), artists ([Alper and Wassall, 2006](#)), movie directors ([John et al., 2017](#)), and actors ([Han and Ravid, 2018](#)), but also in the laboratory ([Casson et al., 2010](#)). However, whereas this literature focuses on the outcomes of participants in winner-take-all markets, we analyze how the presence of superstars influences market *participation* in the first place. These findings also relate to the ongoing discussion about the size and structure of superstar earnings, in particular in sports.

Second, to the best of our knowledge we are the first to incorporate the role of skewness in the earnings distribution in a model of labor supply. This approach connects the “jackpot” gambling literature with the scholarship of human capital investment decisions. Our model is similar to that of [Forrest et al. \(2002\)](#), who find that the size of a jackpot increases lotto demand, even when expected rewards are held constant. However, whereas lotteries entail small stakes gambles with purely random outcomes, we find similar dynamics in high stakes career decisions influenced by individual abilities: The presence of superstars appears to lure more participants into choosing occupations with highly skewed earnings. Our analysis of human capital investments and sequential learning about abilities relates to existing research regarding

(i) whether or not to attend college (Carrell and Sacerdote, 2017), (ii) whether to continue or to dropout once in college (Stange, 2012; Stinebrickner and Stinebrickner, 2012), (iii) the choice of college major (Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2013), and (iv) the returns to college (Hendricks and Leukhina, 2018). Note two key differences, though: We study skill-specific human capital investments for superstar markets, whereas college education primarily entails general human capital development for non-superstar markets. The role of skewness is something our paper adds to the narrative on winner-take-all markets, both from a theoretical and an empirical perspective.

Third and final, the longitudinal pipeline dataset on professional and aspiring professional tennis players we construct constitutes a methodological contribution to the study of winner-take-all markets. This dataset affords three major improvements over other settings in which superstars are observed: our database includes a comprehensive list of virtually all potential entrants into the arena of professional tennis, both women and men, avoiding problems of sample selection and survivorship bias (Denrell, 2003; Jehiel, 2018). The longitudinal nature of the data allows us to identify the effects of skewness from changes over time in a person's predicted earnings distribution and to control for person- and time-specific factors that influence preferences. Finally, the availability of global rankings each year means that objective measures of success are observed by each person throughout his/her formative years. Beyond our analysis presented in the following pages, we hope that researchers will find this database insightful to further explore associated research questions.

2 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_{it} . 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 (also see [Anderberg and Andersson, 2003](#), for educational investment decisions under wage uncertainty).

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:

$$E_t[u(w_i)] + \gamma_i + \epsilon_{it} > u(\hat{w}_i|T - t), \quad (1)$$

where u represents the person's utility function over lifetime income, γ constitutes a person-specific term reflecting individual levels of over- or under-confidence, as well as preferences for the risky career, and ϵ captures the person's calculation or information errors at any point in time. 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](#)):

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

From here, taking expectations given the person's rankings history at time t produces:

$$E_t[u(w_i)] = u(\bar{w}) + u'(\bar{w})E_t(w_i - \bar{w}|\mathbf{r}_{it}) + \frac{u''(\bar{w})}{2}E_t[(w_i - \bar{w})^2|\mathbf{r}_{it}] + \frac{u'''(\bar{w})}{6}E_t[(w_i - \bar{w})^3|\mathbf{r}_{it}], \quad (3)$$

where \mathbf{r}_{it} denotes the sequence of person i 's rankings up to period t . The person will continue

in the risky career after period t if:

$$\epsilon_{it} > u(\hat{w}_i|T-t) - u(\bar{w}) - u'(\bar{w})E_t(w_i - \bar{w}|\mathbf{r}_{it}) - \frac{u''(\bar{w})}{2}E_t[(w_i - \bar{w})^2|\mathbf{r}_{it}] - \frac{u'''(\bar{w})}{6}E_t[(w_i - \bar{w})^3|\mathbf{r}_{it}] - \gamma_i \quad (4)$$

For simplicity, let $\epsilon_{it} \sim U[-e, e]$; then, the probability that the person will continue in the risky career becomes:⁵

$$P(\text{risky career}_{it}) = \frac{e - u(\hat{w}_i|T-t) + u(\bar{w})}{2e} + \frac{u'(\bar{w})}{2e}E_t(w_i - \bar{w}|\mathbf{r}_{it}) + \frac{u''(\bar{w})}{4e}E_t[(w_i - \bar{w})^2|\mathbf{r}_{it}] + \frac{u'''(\bar{w})}{12e}E_t[(w_i - \bar{w})^3|\mathbf{r}_{it}] + \gamma_i. \quad (5)$$

Equation 5 will form the basis of our empirical estimation. It implies that 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. If the person is risk neutral, $u'' = 0$, and if the person is skewness neutral, $u''' = 0$. However, a significant positive coefficient on the cubed term indicates that the person is skewness-loving.⁶

3 Data

3.1 Tennis as a Laboratory to Study Superstar 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) who compete in numerous tournaments throughout Europe (for 2018 alone, there are 418 junior

⁵The assumption of a uniform distribution is not essential but it eases the exposition here.

⁶King (1974), Johnson (1977), and Hartog and Vijverberg (2007) also measure variance and skewness with the standard deviation and the third moment of earnings.

tournaments). The results of these tournaments produce the earliest and most comprehensive global rankings for tennis players. Tennis Europe forms the largest regional federation of the ITF and today manages around 1,200 international tennis events per year. Although European players form the majority on this tour, competitors from 175 countries participate throughout our database (165 origin countries in the male database and 156 origin countries in the female database). Nevertheless, all our results are consistent when focusing on European players only. 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. As with Tennis Europe, the ITF rankings include each player's full name, birthday, nationality, and ranking. Third, the complete data regarding players' professional performances 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 youths in response to objective ranking information.

Crucially, this data set contains a comprehensive pool of potential market entrants: Even those who only play once in any junior competition and lose in their first match 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. Decision-makers under uncertainty cannot learn how market returns and risks are jointly distributed because they only observe idiosyncratic market signals about decisions 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](#)).⁸

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). 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 male and 15,078 female individual player-year observations from 7,242 male and 6,205 female players. Of those who play professional tennis, the average male and female player in this database 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 1](#) plots annual prize money against players’ year-end rankings for 1997, illustrating how professional ranking and professional earnings are closely related. Because prize money is allocated according to the round in which a player loses in a tournament and declines sharply with tournament prestige, the top players

⁷For 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](#)).

⁸For example, by eliminating poorly performing products, mutual funds overstate their performance and understate their risk ([Elton et al., 1996](#)).

account for the vast majority of total prize money earned. Figure 1 documents the extremely unequal distribution of earnings in tennis (the Gini coefficients are 0.87 and 0.68 for women and men, respectively), which is more pronounced than in any other major sport (also see [Beaton, 2014](#)).

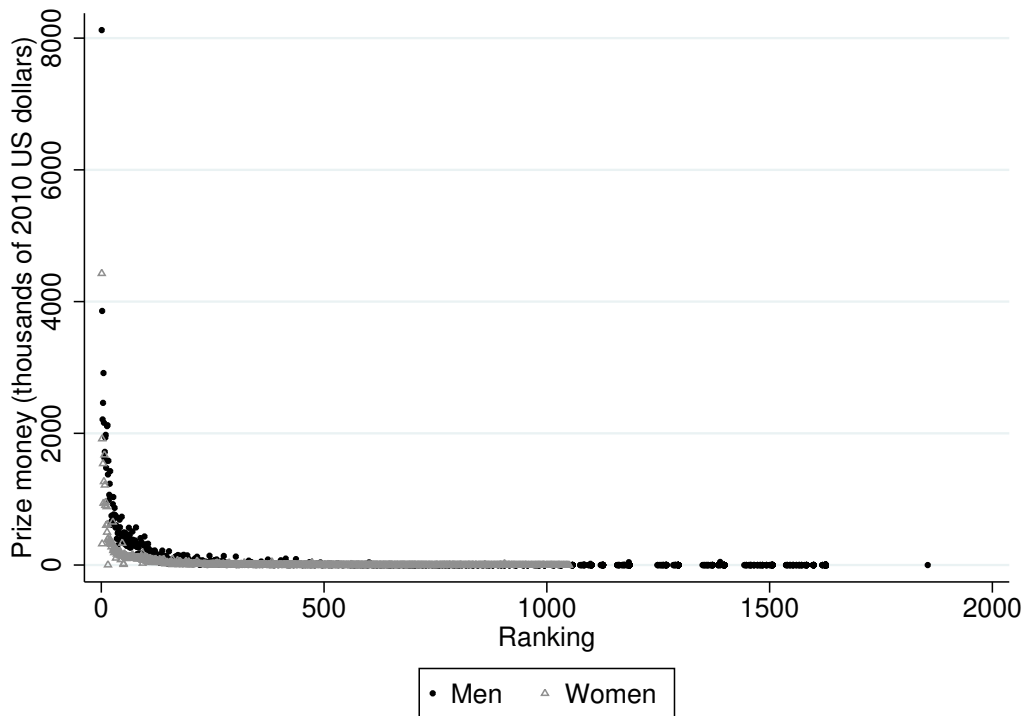


Figure 1: Average annual prize money earnings on the professional Tours in 1997 (y-axis) plotted against professional rankings (x-axis).

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 closely match their expected earnings. However, Figure 2 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 (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 stands in contrast to Rosen’s (1986, p.134) argument, in his review of *The Winner-Take-All Society* (Frank and Cook, 2010; also see Cason et al., 2010, for experimental evidence on winner-take-all markets being inefficient), 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.”

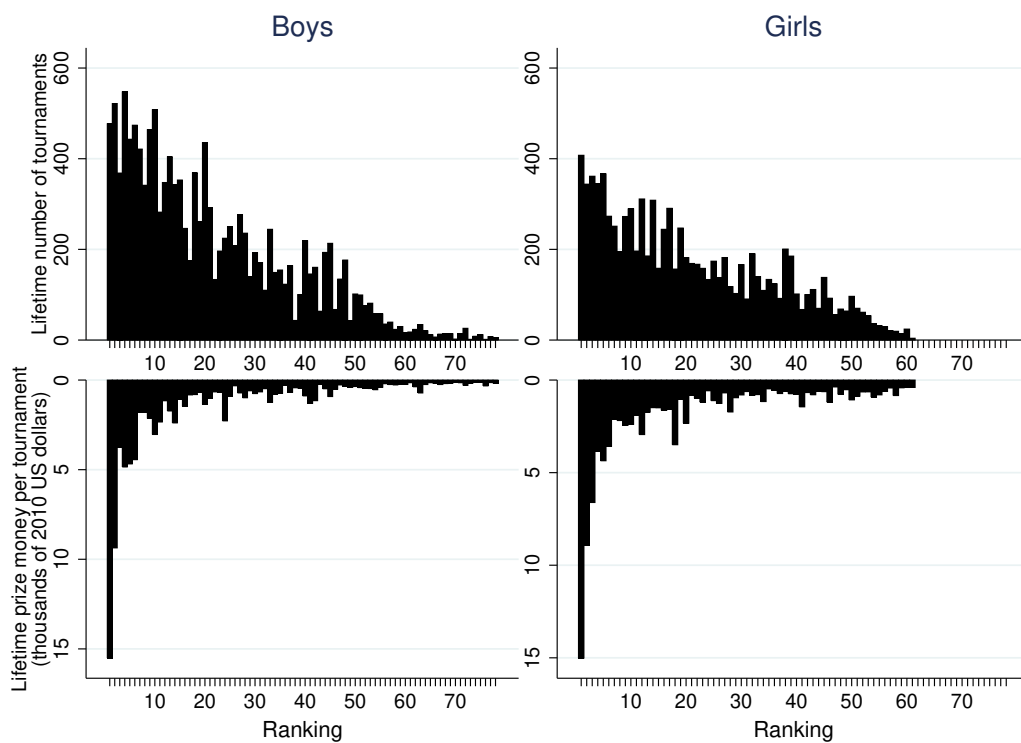


Figure 2: Plotting junior rankings (x-axis) against the lifetime number of professional tournaments played (y-axis in top panel) and lifetime prize money earnings per tournament (y-axis in bottom panel).

3.2 Data Preparation and Descriptive Statistics

3.2.1 Individual Probability Density Functions (PDFs) of Lifetime Prize Money Distributions

Our analysis requires estimates of the lifetime prize money distribution a given player can expect to face at any given age. 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 ranked in more than one age category in a year (e.g., a 16 year-old listed 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.⁹

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 visualized in Figure 2. In that case:

$$W_{it} = W(r_i) = \sum_{v=1}^{b+30-t} w(r_{i(t+v)}). \quad (6)$$

To simplify our analysis, our empirical estimations assume that players only receive prize money between the ages of 19 and 30.¹⁰ Since players do not know the future prize money distribution, we assume they use the distribution of prize money on the professional tour (i.e., WTA and ATP) in the previous year as a reference point.¹¹ Therefore, prize money v years in

⁹In 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.

¹⁰It is noteworthy to point out that professional tennis careers only last until an individuals’ thirties usually, which differs from most conventional careers, in which people work 20-30 years more. To the extent that such particularities of professional tennis are common knowledge and remained by and large fixed throughout our sample period, this should not affect our empirical analysis.

¹¹Young players might realistically assume a certain amount of growth in prize money, but as long as they do

the future is given by:

$$w(r_{i(t+v)}) = w(r_{t+v}, a_{it} + v, s_i, t - 1), \quad (7)$$

where the right-hand side captures 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 3rd among his birth cohort in 10 years' time would expect to earn the same amount (adjusted for inflation) as the 3rd best 23 year-old man in the professional tour in 1991. To simplify computations, prize money is rounded to the nearest \$1,000.

The probability density function (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:

$$P(W_{it}) = P(r_{i(t+30-a)}, r_{i(t+29-a)}, \dots, r_{i(t+1)}). \quad (8)$$

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-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:

$$P(W_{it}) = P(r_{t+v}|r_{t+v-1}, a_{it} + v - 1, s_i)P(r_{t+v-1}|r_{t+v-2}, a_{it} + v - 2, s_i) \dots P(r_{t+1}|r_{it}, a_{it}, s_i). \quad (9)$$

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

not anticipate a change in the skewness of the distribution, our key results will be unaffected.

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.

We then calculate lifetime prize money PDFs for each active player in every year. The estimated lifetime prize money PDFs exhibit significant variation. 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 derived lifetime PDFs exhibit significant variation. As an example, Figure 3 plots the distributions faced by 18-year olds in 1997 (the midpoint of our database) 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 Fisher-Pearson coefficients of skewness for the distributions (calculated as $\frac{\frac{1}{n} \sum (W_i - \bar{W})^3}{\left(\frac{1}{n} \sum (W_i - \bar{W})^2\right)^{\frac{3}{2}}}$) 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 4 and 5 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 face 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.

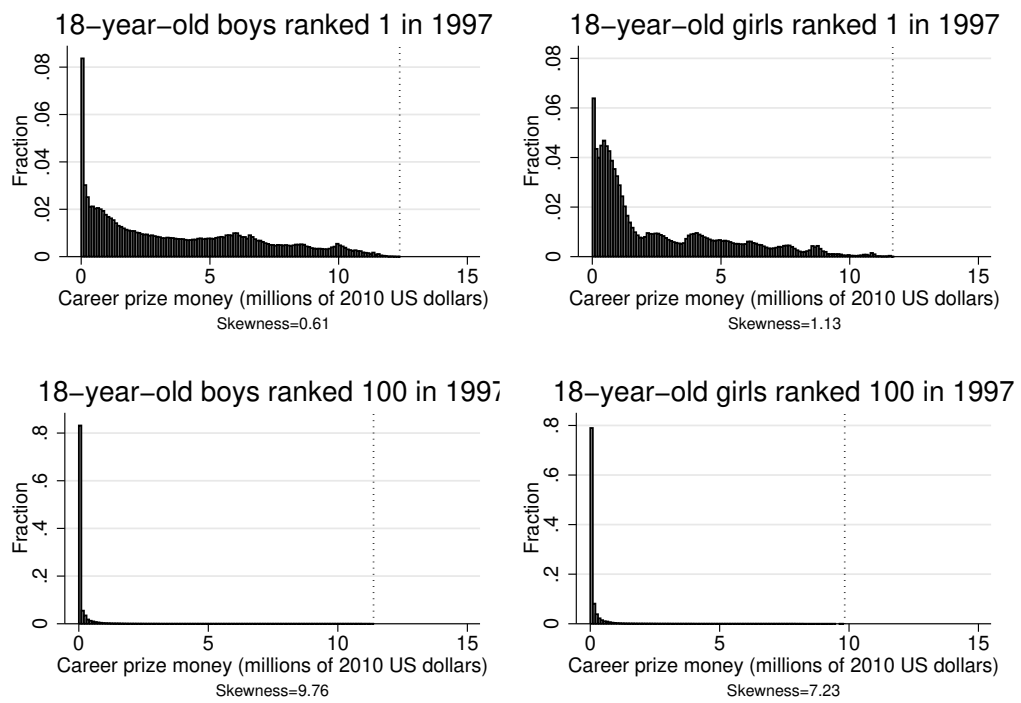


Figure 3: Plotting the expected distribution of career prize money for #1's (top panel) and #100's in the U18 rankings.

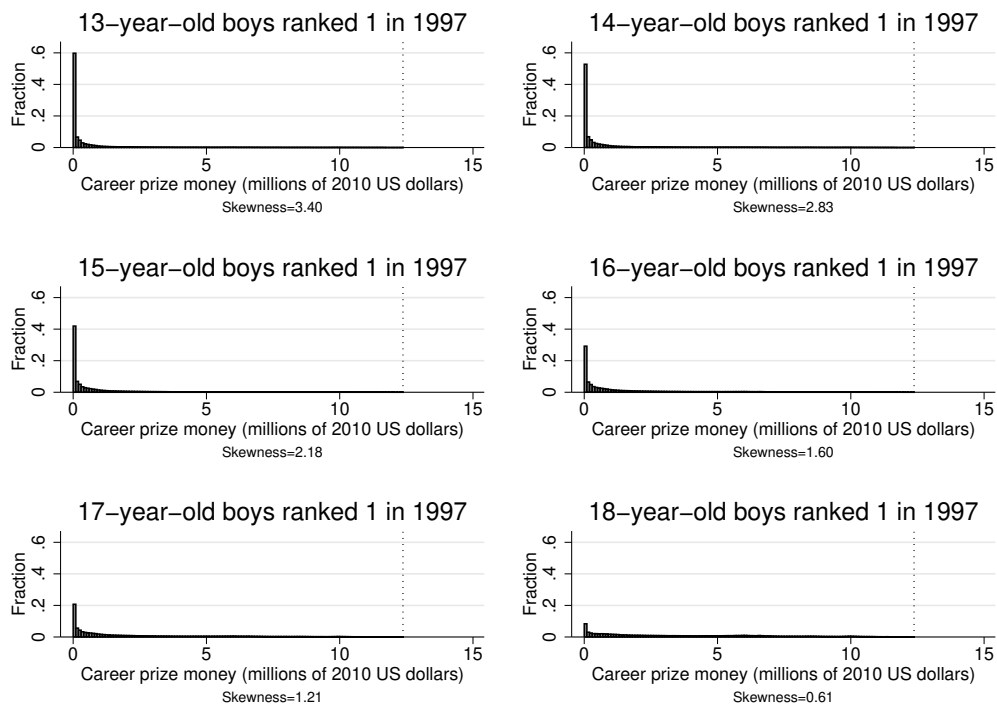


Figure 4: Plotting the expected distribution of career prize money for male #1's at the ages 13 until 18.

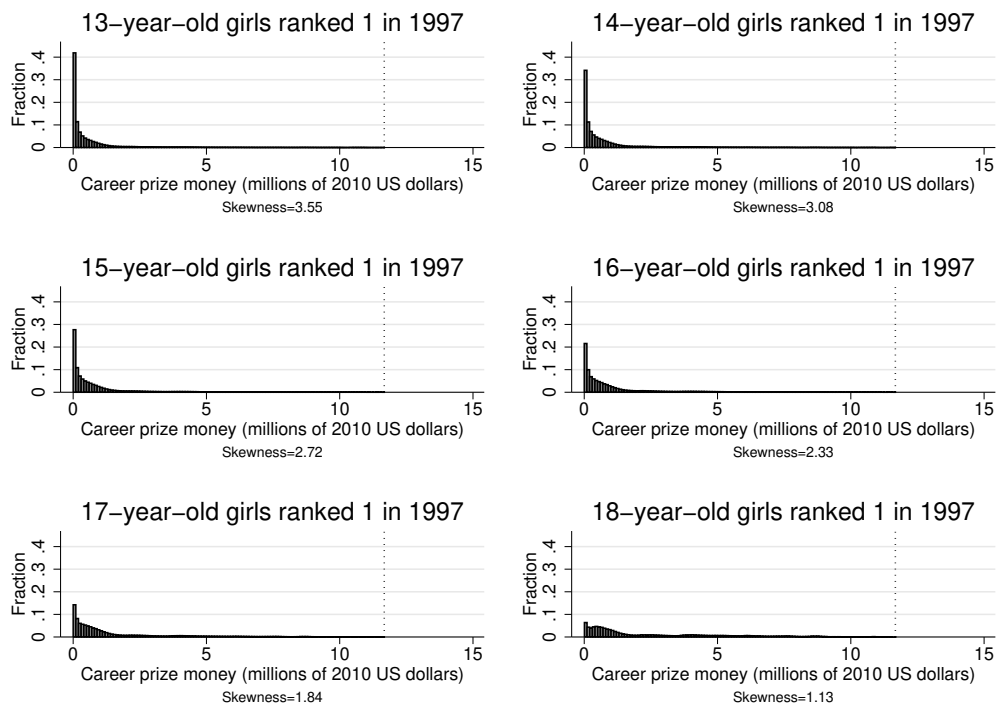


Figure 5: Plotting the expected distribution of career prize money for male #1's at the ages 13 until 18.

3.2.2 Individual Mean, Variance, and Skewness of Prize Money

We can calculate three moments of the lifetime earnings distributions for any player i in year t , analogous to the three moments from equation (5):

$$\begin{aligned} E(W_{it}) &= \sum_W P(W_{it})W_{it}; \\ E[W_{it} - E(W_{it})]^2 &= \sum_W P(W_{it})[W_{it} - E(W_{it})]^2; \\ E[W_{it} - E(W_{it})]^3 &= \sum_W P(W_{it})[W_{it} - E(W_{it})]^3. \end{aligned} \quad (10)$$

Figure 6 plots the corresponding values of these three moments for each age between 13 and 29, separately by gender and for rankings 1, 10, and 50. Average predicted prize money is highest for higher-ranked men and women but falls rapidly between ranks 1 and 10. The mean rises up to age 19 for women and 20 for men who are ranked 1 or 10; however, it falls during this period for those who are ranked 50. After age 20, the mean falls as the number of years left in a player's career shrinks. A similar pattern is found for variance and skewness. Men have higher values of all three moments than women, for any age and ranking.

3.2.3 Labor Supply in Tennis

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). Nevertheless, results are consistent when raising that bar to defining 'active' on the professional Tours as having played, for example, seven tournaments per year. 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 face a higher chance of earning very

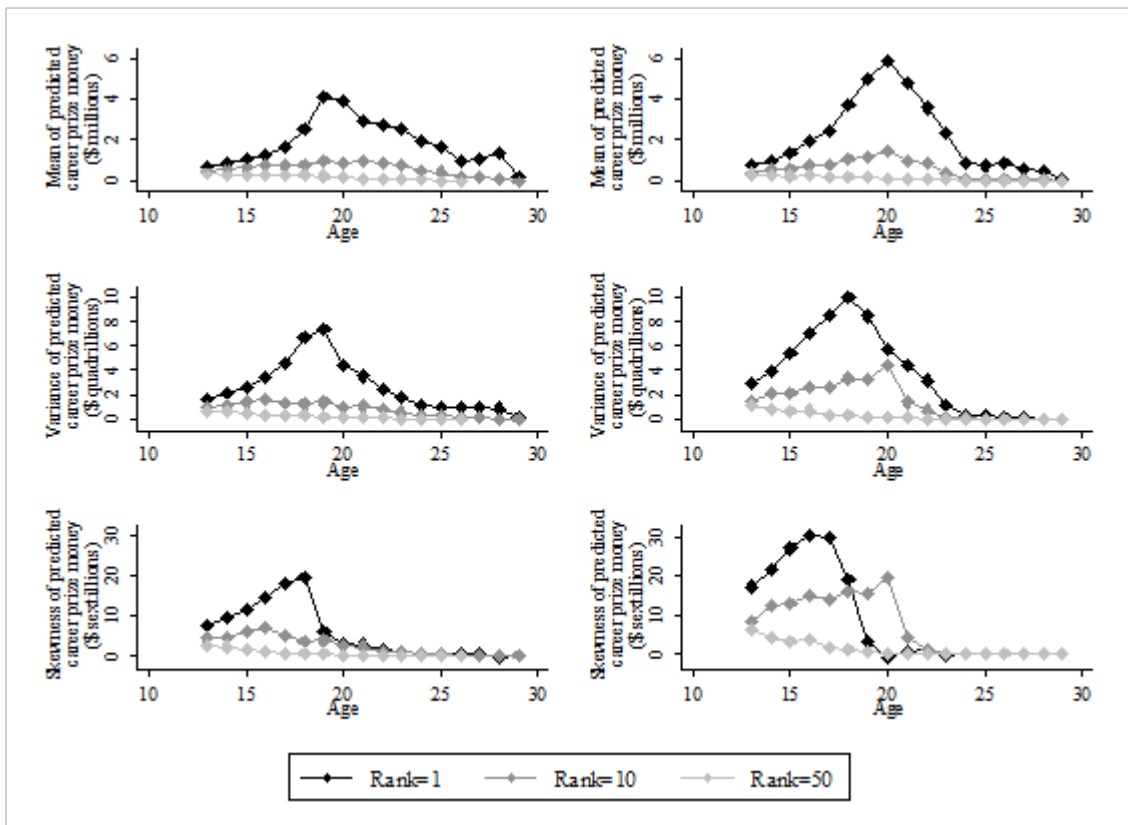


Figure 6: Moments of the predicted prize money distributions in 1997 by age, rank, and gender. Results for girls are shown on the left, whereas results for boys are shown on the right.

little over their careers. Male tennis players also face a much greater variance and skewness than female tennis players in their expected earnings.

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 (0.497)	0.657 (0.475)	-0.105	0.000***
Mean of predicted career prize money (millions of 2010 US\$)	0.290 (0.353)	0.298 (0.251)	-0.008	0.030**
Variance of predicted career prize money (trillions of 2010 US\$ squared)	0.899 (1.634)	0.499 (0.837)	0.400	0.000***
Skewness of predicted career prize money (in quintillions of 2010 US\$ cubed)	5.653 (10.949)	2.143 (4.584)	3.510	0.000***
Lagged ranking	119.150 (91.358)	87.029 (64.982)	32.121	0.000***
Age	16.862 (1.509)	16.679 (1.519)	0.183	0.000***
Players	7,242	6,205		
Number of observations	15,810	15,078		

Notes: *, **, and *** denote the means for males and females to be significantly different from each other at the 10%, 5%, and 1% level, respectively.

The fact that lagged rankings are worse on average (i.e., take higher values) for males than for females reflects the fact that more males are competing in each age group than females, on average. Table 2 reports how many junior players became internationally ranked at each age and for how many years they continued to be internationally ranked. Players exhibit both steady attrition and yet persistence many years later. For example, of those first ranked at age 14, 13 percent of the females and 15 percent of the males exited after a year, and yet 23 percent of the females and 25 percent of the males remained globally ranked six years later.

Table 2: Number of observations per player aged 14-19.

Age of first ranking appearance	Number of years in junior rankings						Total
	1	2	3	4	5	6	
Males (N = 7,242)							
14	119	127	135	192	151	218	942
15	597	463	435	295	264		2,054
16	315	273	145	159			892
17	993	374	388				1,755
18	455	437					892
19	707						707
Females (N = 6,205)							
14	204	177	214	248	149	338	1,330
15	380	317	292	205	336		1,530
16	400	382	170	216			1,168
17	702	232	256				1,190
18	369	264					633
19	354						354

4 Empirical Findings

4.1 Empirical Strategy

In our most complete estimation, 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 $t - 1$, consistent with equation (5), as follows:¹²

$$Active_{it} = \alpha_1 E_t(W_i) + \alpha_2 E_t(W_i - \bar{W})^2 + \alpha_3 E_t(W_i - \bar{W})^3 + \gamma_i + \mathbf{Birthyear}_i \times \mathbf{Age}_{it} \beta + \epsilon_{it}, \quad (11)$$

where γ_i introduces a player fixed effect, controlling for differences in innate preferences for tennis, over-and under-confidence, and other person-specific factors (as noted in Section 2). $\mathbf{Birthyear}_i$ represents a full set of birth cohort dummies and \mathbf{Age}_{it} constitutes a full set of age dummies. The interaction of these terms captures the effects of changes in opportunity cost over a person's teenage years, as well as observed changes in the distribution of professional earnings that affect all junior players in the same cohort and year equally. Finally, ϵ_{it} constitutes a random error term, and we cluster standard errors at the player level to allow for the fact that errors might be correlated over time for each player. Throughout our analysis, we analyze males and females separately, using linear probability models.

4.2 Main Findings

Tables 3 and 4 report our main regression results for males and females. In each Table, columns (1) – (6) analyze player-year observations for ages 14-19, whereas column (7) investigates those aged 20-28. To begin with, we regress the participation dummy variable on the mean, variance, and skewness of an individual player's expected earnings only. As reported in column (1) of

¹²Our results are consistent when estimating a logit model with random effects.

both Tables, all three coefficients are statistically significant at the one percent for males and females, exhibiting the signs we anticipate based on theory: higher means and skewnesses increase the likelihood of remaining in tennis, whereas higher variances decrease it. In column (2), we add age and player fixed effects, but the signs and statistical relevance of all coefficients of interest remain unchanged.

Column (3) reports our main results, incorporating player and cohort-age fixed effects, as specified in equation 11. Signs and statistical significance remain consistent with the estimates displayed in columns (1) and (2), both for males and females. In terms of magnitude, the estimated coefficients for males imply an elasticity with respect to the mean of prize money of 0.574 (equivalent to multiplying the derived coefficient with the average mean earnings and dividing by the average participation rate). Similarly, the implied elasticities with respect to males' variance and skewness in expected prize money are -0.471 and 0.184. We can conduct a simple back-of-the-envelope calculation to illustrate the corresponding magnitude: in the male tennis player sample, skewness ranges from 0.50 to 17.88, with a mean of 7.57. Now, if the average Fisher-Pearson coefficient of skewness in our data were reduced to the level of the entire US labor market in the 2010 American Community Survey (which is 3.99), holding constant means and variances, the average male player would be 7.2 percent less likely to continue playing tennis the following year.

For females, the coefficients in column (3) imply elasticities with respect to the mean, variance, and skewness of prize money equal to 0.320, -0.197, and 0.075, respectively. The estimated coefficient on the mean is significantly lower than the coefficient derived for males, but the coefficients on the variance and skewness are not significantly different between genders. However, females have a considerably higher average participation rate and lower average skewness compared to males (as shown in Table 1). Therefore, the elasticity of participation with respect to skewness for females is significantly lower than the elasticity for males. If the Fisher-Pearson coefficient of skewness were set equal to the 2010 US labor market value,

Table 3: Regression results from estimating the participation equation in year t for males.

	Ages 14-19				Ages 20-28		
	(1)	(2)	(3)	(4)	(5) Only Europeans	(6) Costs adjusted	(7)
Mean career prize money (millions of 2010 US\$) $_{t-1}$	1.016*** (0.065)	1.162*** (0.093)	1.030*** (0.091)		0.860*** (0.107)	1.094*** (0.107)	0.120*** (0.025)
Variance career prize money (trillions of 2010 US\$ 2) $_{t-1}$	-0.302*** (0.023)	-0.341*** (0.031)	-0.289*** (0.031)		-0.245*** (0.036)	-0.302*** (0.035)	-0.056*** (0.013)
Skewness career prize money (quintillions of 2010 US\$ 3) $_{t-1}$	0.025*** (0.002)	0.022*** (0.002)	0.018*** (0.002)		0.016*** (0.003)	0.019*** (0.003)	0.003** (0.002)
Mean career prize money \times Aged 14/15				2.018** (0.833)			
Mean career prize money \times Aged 16/17				0.646*** (0.251)			
Mean career prize money \times Aged 18/19				1.055*** (0.124)			
Variance career prize money \times Aged 14/15				-0.510 (0.341)			
Variance career prize money \times Aged 16/17				-0.191** (0.095)			
Variance career prize money \times Aged 18/19				-0.297*** (0.039)			
Skewness career prize money \times Aged 14/15				0.022 (0.026)			
Skewness career prize money \times Aged 16/17				0.014* (0.007)			
Skewness career prize money \times Aged 18/19				0.018*** (0.003)			
[0.5em] Age fixed effects		✓					
Player fixed effects		✓	✓	✓	✓	✓	✓
Birthyear \times age fixed effects			✓	✓	✓	✓	✓
R ²							
Number of observations	0.071 15,810	0.618 15,810	0.649 15,810	0.649 15,810	0.633 10,581	0.647 15,810	0.616 6,009

Notes: In column (5), prize money is adjusted for the estimated costs of competing. Standard errors are clustered at the player level and presented in parentheses. *, ** and *** denote significance at the 10 percent, 5 percent, and 1 percent level, respectively.

Table 4: Regression results from estimating the participation equation in year t for females.

	Ages 14-19			Ages 20-28			
	(1)	(2)	(3)	(4)	(5) Only Europeans	(6) Costs adjusted	(7)
Mean career prize money (millions of 2010 US\$) $_{t-1}$	1.559*** (0.057)	0.608*** (0.066)	0.706*** (0.067)		0.745*** (0.078)	0.787*** (0.108)	0.247*** (0.044)
Variance career prize money (trillions of 2010 US\$ 2) $_{t-1}$	-0.510*** (0.033)	-0.192*** (0.033)	-0.260*** (0.034)		-0.270*** (0.038)	-0.339*** (0.055)	-0.106*** (0.023)
Skewness career prize money (quintillions of 2010 US\$ 3) $_{t-1}$	0.030*** (0.004)	0.014*** (0.004)	0.023*** (0.004)		0.023*** (0.004)	0.038*** (0.006)	0.001 (0.004)
Mean career prize money \times Aged 14/15				-0.043 (0.578)			
Mean career prize money \times Aged 16/17				-0.407** (0.167)			
Mean career prize money \times Aged 18/19				0.770*** (0.098)			
Variance career prize money \times Aged 14/15				0.196 (0.570)			
Variance career prize money \times Aged 16/17				-0.218** (0.111)			
Variance career prize money \times Aged 18/19				-0.278*** (0.049)			
Skewness career prize money \times Aged 14/15				-0.041 (0.077)			
Skewness career prize money \times Aged 16/17				0.030** (0.014)			
Skewness career prize money \times Aged 18/19				0.022*** (0.005)			
Age fixed effects		✓					
Player fixed effects		✓	✓	✓	✓	✓	✓
Birthyear \times age fixed effects			✓	✓	✓	✓	✓
R ²	0.079	0.643	0.659	0.660	0.649	0.657	0.561
Number of observations	15,078	15,078	15,078	15,078	9,988	15,078	4,693

Notes: In column (5), prize money is adjusted for the estimated costs of competing. Standard errors are clustered at the player level and presented in parentheses. *, ** and *** denote significance at the 10 percent, 5 percent, and 1 percent level, respectively.

females would be only 2.6 percent less likely to stay in tennis, on average.

In column (4), we examine the variation in the effects of the three moments across age by interacting the mean, variance, and skewness variables with binary indicators for whether a person is aged 14-15, 16-17, or 18-19. These estimations aim to delineate whether the relationships we derive vary across particular age groups. What we find are notable age differences. First, while mean earnings constitute a positive predictor of participation for males across all age groups, this is only the case for females aged 18/19 – not for younger girls. In fact, higher expected prize money emerges as a *negative* predictor of participation for female players aged 16/17. Second, variance and skewness do not drive participation of 14/15 year-old male or female players. Thus, the corresponding results from column (3) are largely driven by those aged between 16 and 19.

4.3 Robustness Checks

Beyond our main results, columns (5) – (7) of Tables 3 and 4 display findings from several additional specifications. First, since all U14 and U16 tournaments are contested in Europe, it is possible that including non-European players 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 column (5). 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 0.018 to 0.016, whereas that for females remains consistent at 0.023.

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, since top players usually have these costs paid for by sponsors or tournament organizers, we may understate the underlying skewness in earnings.¹³ 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.¹⁴ The corresponding results are displayed in column (6). The skewness elasticity remains consistent for males with a coefficient of 0.019 and further rises for females to 0.038.

Third and final, column (7) displays results when the database is changed to ages 20-28, rather than ages 14-19.¹⁵ 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 coefficients for both male and female players, with coefficients being cut to 0.003 for males and 0.001 for females. Whereas the relationship for males remains statistically significant at the five percent level, that for females becomes indistinguishable from zero. Thus, tennis players appear much less influenced by the skewness of income once they have entered the professional tennis labor market than when they are aspiring to enter in their teenage years. These results suggest that winner-take-all markets lure young entrants into seeking the equivalent of ‘lottery wins’ in their careers, at least in the labor market for professional tennis.

¹³In fact, at the highest-level professional events, accommodation is always complementary until a player loses in the respective tournament. Further, top players frequently receive appearance fees at most non-Grand Slam events.

¹⁴Note 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.

¹⁵Note 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.

5 Conclusion

The Winner-Take-All Society (2010), [Frank and Cook](#)'s best-selling book that has been cited more than 2,600 times in 2019, turned Adam Smith's "career lotto" conjecture into a global conversation. But even before that, economists critiqued both the lack of a theoretical explanation for labor supply decisions in winner-take-all markets, as well as the absence of comprehensive data with which to analyze such decisions ([Rosen, 1986](#); [Galbraith, 1995](#)). This paper provides some of the first empirical evidence of whether and how labor supply in a winner-take-all market is systematically driven by the earnings of the very best – independent of the mean and variance of one's own expected earnings.

We aim to make two contributions. First, we formalize 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 the mean and variance of expected earnings. Second, we overcome sample selection and survivorship bias problems by assembling a longitudinal pipeline dataset of a comprehensive pool of potential market entrants into professional tennis – a typical winner-take-all market ([Denrell, 2003](#); [Jehiel, 2018](#)). 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 player's lifetime earnings in tennis.

In our benchmark estimation, we find skewness in the individual player's expected prize money distribution to account for 7.2 percent of the male teenage players who continue to pursue professional tennis careers. Similarly, 2.6 percent of female players continue with their tennis careers because of the skewness in their individualized expected earnings. These findings are consistent with the hypothesis that the presence of superstars encourages modestly-talented people to pursue longshot career gambles. In further estimations, we find that skewness plays a particularly pronounced role for players aged between 16 and 19. All results are derived when

controlling for the mean and variance of expected earnings, as well as player- and cohort-age fixed effects. Accounting for player-fixed effects controls for personal preferences about career choices (e.g., an individual's enjoyment of playing tennis) that remain constant over an individual player's teenage years. Nevertheless, we may fail to capture such personal preferences if they systematically change throughout a player's teenage years.

What can explain this skewness finding? We can think of two possible narratives: a general preference for skewness and overconfidence. The latter hypothesis would suggest that moderately talented players 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 an enormous amount of money may strengthen a player's decision to continue with a tennis career, precisely because he is overestimating the likelihood of being the next Roger Federer.

However, by controlling for player- and cohort-age fixed effects in our specifications, we account for those effects from overconfidence that do not change over time for any individual between the ages of 14 and 19, as well as overconfidence that is typical for specific cohorts at certain ages. Hence, for overconfidence to fully explain the patterns observed in our results, it would have to be correlated with skewness for each player over time, meaning that a given player becomes more overconfident as he/she falls down the rankings from year to year (and vice versa). This seems unlikely, suggesting that skewness preferences are an important factor to explain our results. We hope that future research will be able to further distinguish between these two hypotheses. More generally, we hope this paper provides an empirical starting point for our understanding of entry into superstar markets.

References

- Alper, N. O. and Wassall, G. H. (2006). Artists' careers and their labor markets. *Handbook of the Economics of Art and Culture*, 1:813–864.
- 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.
- Anderberg, D. and Andersson, F. (2003). Investments in human capital, wage uncertainty, and public policy. *Journal of Public Economics*, 87(7-8):1521–1537.
- 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 jahansson lectures). *Yrjo Jahanssonin Saatio, Helsinki*.
- 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.
- 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.
- 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.
- 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.
- Carrell, S. and Sacerdote, B. (2017). Why do college-going interventions work? *American Economic Journal: Applied Economics*, 9(3):124–51.
- Cason, T. N., Masters, W. A., and Sheremeta, R. M. (2010). Entry into winner-take-all and proportional-prize contests: An experimental study. *Journal of Public Economics*, 94(9-10):604–611.

- 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*.
- Cook, P. J. and Clotfelter, C. T. (1993). The peculiar scale economies of lotto. *The American Economic Review*, 83:634–643.
- 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. (2020). Saliency and skewness preferences. *Journal of the European Economic Association*, 18(5):2057–2107.
- 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.
- 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.
- 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.
- Gloster, R. (2015). The brutal costs of raising the world's 631st-best tennis player. Bloomberg Business. Accessed: 2015-09-21.
- Golec, J. and Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *Journal of Political Economy*, 106(1):205–225.
- Han, S. and Ravid, S. A. (2018). Star turnover and the value of human capital: Evidence from Broadway show. *Management Science*, *Forthcoming*.
- Hartog, J. and Vijverberg, W. P. (2007). On compensation for risk aversion and skewness affection in wages. *Labour Economics*, 14(6):938–956.

- Hendricks, L. and Leukhina, O. (2018). The return to college: Selection and dropout risk. *International Economic Review*, 59(3):1077–1102.
- 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.
- John, K., Ravid, S. A., and Sunder, J. (2017). Managerial ability and success: Evidence from the career paths of film directors. *Journal of Corporate Finance*, 44:425–439.
- 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.
- 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.
- 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.
- 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.
- MacDonald, G. (1988). The economics of rising stars. *American Economic Review*, 78(1):155–66.
- 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.
- 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.
- Rossingh, D. (2011). Next Sharapovas cost \$400,000 as teen champs disappear. Bloomberg Business. Accessed: 2015-09-21.
- Rossingh, D. (2016). Want your kid to become a tennis player? Prepare to spend big. CNN. Accessed: 2015-09-21.
- 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.