Career Lotto? Labor Supply in Winner-Take-All Markets

Wayne A. Grove, Michael Jetter and Kerry L. Papps

How does the presence of superstars, who earn vast amounts, affect the participation of more modestly-talented workers in a winner-take-all labor market? We use longitudinal data on junior and professional tennis players to examine how the decision to pursue this extremely risky career is related to the earnings a player is likely to receive in the future. We find that teenage players prefer high mean earnings and a low variance, but are also attracted to highly-skewed earnings distributions – just like gamblers at horse races or in lotteries. The magnitudes of the skewness effects we estimate are much smaller than in studies of regular gambling behavior; however, they are still sizeable. If the skewness were to fall to zero, boys would be around 20% less likely on average to continue in the sport, whereas girls would be 5% less likely to continue. A side-effect of winner-take-all labor markets therefore is an increase in participation, even among those who have a negligible chance of earning large amounts.

1. Introduction

> "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)

"Never give up your dreams" is the inspirational message uttered by Oscar winners, Olympic medalists, award-winning authors, and others in arts, culture, entertainment, and sports markets that offer miniscule probabilities of truly life-changing outcomes.[1] Does the presence of those and other superstars, who enjoy fame and fortune, cause substantial individual and social inefficiency by encouraging modestly-talented workers to pursue long-shot careers with negative expected values, rather than more realistic occupational paths?[2] Doubting that such aspirants act as "giddy risk lovers with unrealistic assessments of themselves", Rosen and Sanderson (2001) speculate that continuous feedback on one's performance causes entrants to switch to more realistic

[1] For example, the 2017 female winner of the New York City marathon, Shalane Flanagan, the first American in 40 years, said: "I've been dreaming of a moment like this since I was little girl… Hopefully it inspires the next generation of women… It took me seven years to do this" (Schonbrun, 2017).

[2] Frank and Cook (1995), in *The Winner-Take-All Society,* most prominently advance the idea that the "winner-take-all payoff structure [in superstar markets] generates a spiral of individual and social occupational waste, since it leads both to increasing (monetary and non-monetary) reward inequalities and to overcrowdings in the markets and occupations prone to an overestimation of one's chance to succeed". For example, the documentary film *Hoop Dreams* depicts the ultimately unsuccessful quest of two African-American teenagers from poor Chicago neighborhoods to play for an NBA team. For example, a quarter of the parents whose children merely participate in high school sports hope their children will play in the pros, with much higher percentages for less educated and less affluent parents (NPR 2015, p.17).

careers when the prospects of making the big leagues gets sufficiently unfavorable. Ample evidence, though, contradicts such expected utility theory predictions that individuals make decisions under uncertainty based on the probabilities of outcomes.[3] For example, the size of the jackpot, not the odds of winning appears to drive lottery ticket sales.[4] Hence, we address the puzzle of why so many people play "career lotto" by estimating whether superstars' success influences potential entrants to make large financial investments in exceedingly risky, long-shot occupations.[5]

Alfred Marshall's (1920) observation almost a century ago, prior to the diffusion of broadcast and recording technologies, that singers' salaries were constrained by the limited "number of persons who can be reached by a human voice" anticipated the now ubiquitous superstar or winner-take-all markets; in such markets, a very small number of performers satisfy most customers' demands and earn the vast majority of the income.[6] Since professional ability in those markets only reveals itself by on-the-job talent discovery, to be a contender, early aptitude must be nurtured and developed, which may lead to 10-15 years of skill-specific human capital investments; such costs are estimated to exceed $100,000, for example, for tennis.[7] The National Basketball Association (NBA) illustrates the tiny odds of success in superstar labor markets: some 10,000 impressive high school players (of over 500,000 participants) vie for about 1,000 Division I freshman spots who in turn compete to be one of about 30 to join an NBA roster each year and earn over half a million dollars, of whom only about half remain on the rooster for five years.[8] In the past, some organizations bore talent development risks by signing exclusive long-run labor contracts with a portfolio of

---

[3] See the literature on prospect theory (Kahneman and Tversky (1979) and related salience models (Dertwinkel-Kalt and Koster 2017).

[4] See Clotfelter and Cook (1993) and the extensive field and laboratory literature on gambling.

[5] According to Caves (2000), one in 15,000 submitted fiction manuscripts gets published and then in 1994, for example, four authors (Clancy, Crichton, Grisham and King) accounted for 70 percent of the sales (Sorensen 2006).

[6] Although Marshall (1920) assumed that those "reached by the human voice" would pay for the pleasure to do so, the pirating of recordings and performances due to advances in digitalization undermines artists' ability to receive compensation for their performances.

[7] See USTA and Lawn Tennis Association estimates.

[8] See Rosen and Sanderson, 2001. Reports by the National Collegiate Athletic Association (NCAA) suggest extremely low odds of college athletes playing on professional football, baseball, and hockey teams. Although no comparable data on aspiring orchestra members is provided, Christian Colberg, principal viola of the Cincinnati Symphony, suggests that "from a statistical chance, it's probably easier to get into the NBA".

aspiring performers in, for example, the motion picture "studio system" for Hollywood actors and in professional sports.[9] Since the demise of exclusive long run labor contracts, potential entrants to winner-take-all markets, or some backer, must bear the development and annual opportunity costs for a minute chance of superstardom.

If such career decisions really cause substantial social inefficiency, what market failures are to blame? We investigate two possible reasons for individual miscalculations of their odds of success: (1) the lack of data about the probabilities of earning a living in a domain and (2) decision-making based on the desirability of life-changing outcomes, rather than their probability.[10]

Although entrants to superstar markets are often assumed to be risk-lovers, might their risk-taking really reflect attraction to positive skewness as shown by the considerable empirical literature of gambling on horse races and lotteries? Gamblers who accept a lower expected payoff in return for a small chance of a large gain are attracted to skewness, not to risk.[11] In fact, evidence of the willingness to accept a lower expected payoff in return for greater skewness, rather than being risk loving, seems to describe decision-making in labor markets (Hartog and Vijerberg 2007) and regarding entrepreneurship (Chen *et al*. 2016), savings (Gollier 2001), and financial investments (Brunnermeier and Parker 2005). Since these literatures contain surprisingly limited information about skewness attraction differences by gender, a contributions of this paper is doing so for occupational choice decisions.

What is missing in the superstar and winner-take-all markets literature, and the most basic dilemma for the families and young aspiring performers, is that "data are not available to calculate meaningful success probabilities for potential entrants"

[9] Before free agency, Major League Baseball players signed lifelong contracts with teams and in the 1920s-1940s Hollywood actors signed 7-year contracts with motion picture companies (Terviö 2009).

[10] Although most of the empirical and experimental studies cannot distinguish between preferences for positive skewness versus miscalculations about the probability of such outcomes, Snowberg and Wolfers (2010) attempt to distinguish between preference-based and perception-based explanations and find evidence that misperceptions of probabilities account for the longshot bias among horse track gamblers.

[11] Regarding the role of skewness in lottery and horse race gambling, see Clotfelter and Cook (1993), find lotto sales to be positively related to the size of the jackpot, but negatively to expected value. Golec and Tamarkin (1998), and Garrett and Sobel (1999, p.88) who conclude "that lottery players, like horse race bettors, are risk averse but favor positive skewness". For the voluminous laboratory experimental literature regarding the longshot anomaly and evidence for positive skewness preferences, see Grossman and Eckel (2015).

(Rosen and Sanderson 2001).[12] While young violinists, golfers, and opera singers, for example, may focus on the spectacular success of outlier superstars, the lack of information about the entire pool of unsuccessful contenders prevents systematically calculating the likelihood a performer to earn a living in those markets, given their development trajectory. Since no objective quality measures exist for performers in the artistic, cultural, and entertainment markets,[13] Krueger (2005), for example, calculates rock star quality by counting "the number of millimeters of print columns (including photos) devoted to each artist in *The Rolling Stone Encyclopedia of Rock & Roll*" (p. 18).[14] In his seminal article "The Economics of Superstars", Rosen (1981) asserts that "confidentiality laws and other difficulties make it virtually impossible to obtain systematic [earnings] data in this field," referring to classical musicians, comedians and others in "show business." Many professional sports offer solutions to those problems since earnings and productivity information are publically available for at least some sports.[15]

Beyond professional earnings and productivity data, what has proven especially elusive for researchers has been obtaining pre-professional objective measures of individual ability that relate to adult success in the same domain, combined with professional productivity and earnings data.[16] Consequently, very limited empirical analyses exists of potential entrants probabilities or determinants of professional success, even though the extensive literatures of talent development studies (Baker *et al*. 2017), prodigy studies (Lubinski *et al*. 2014), expertise studies (Ericsson *et al*. 2006), and ability selection in sports and labor economics are premised on the early-in-

[12] Although Rosen and Sanderson (2001) refer to professional sports, the same applies to the arts and other forms of entertainment as Caves (2000) discusses in *Creative Industries*.

[13] See Menger (2006). Although numerous competitions exist in the arts, cultural and entertainment markets, we know of no analysis assessing how predictive the winners of the youth competitions (based on objective judgments) are of career outcomes in those domains. Existing studies that compare youth to adult outcomes typically measure success as the achievement of an extraordinary accomplishment, such as standing on an Olympic podium, becoming a top ranked player, or winning a prestigious competition.

[14] Entertainment is more "about finding out the tastes and whims of the public than about the objective measure of quality" (Tervio 2009, 944).

[15] Hence, Kahn's (2000) article entitled: "The Sports Business as a Labor Market Laboratory." Of course, objective productivity measures of athletes vary by sport, especially for team sports. And total earnings may not be available since endorsements and others forms of income besides salaries or prize winnings are not necessarily publically available.

[16] Krueger's (2005) approach can offer no information about the labor supply decisions of potential entrants. Hamlin (1991 and 1994) rated singer's voice quality.

life observability of talent or of particular characteristics thought to predict adult outcomes.[17]

This paper makes three key contributions. First, we contribute to the literature by assembling the first systematic and representative longitudinal dataset of potential entrants and of professionals in a superstar market by linking objective youth rankings of abilities to adult rankings and earnings in the same domain, from age 13 to 27 (the length of almost all careers in the domain).[18] Importantly, the availability of data for males and females permits us to investigate gender differences.

We study tennis because it is the only domain, to our knowledge[19], which offers global ordinal rankings and ranking points for junior players[20] which we link to professional tennis circuit tournament data and earnings for women and men, gathered from the websites for the Women's Tennis Association (WTA) and for the Association of Tennis Professionals (ATP), respectively. Tennis offers objective ability measures based on individual performance outcomes in tennis matches and tournaments, whereas the subjective judgements of scouts determine the rankings of athletes in team sports[21] and decisions by judges in artistic, cultural, and entertainment performers. This youth/amateur sample reasonably constitutes a global pool of potential entrants to this labor market.

---

[17] Although there are some predictors of youth characteristics that predict general adult outcomes, such as the fact that "the top 1 percent of mathematical reasoning ability" 13-year olds in the U.S. far exceeded average outcomes (Lubinski *et al*. 2014), our interest is about domain-specific outcomes and specifically an activity in which youth develop expertise and consider pursuing a career in that domain.

[18] Most studies of success in winner-take-all markets analyze the determinants of achieving a goal, such as winning a prestigious award, competition, or a spot on an Olympic team or an Olympic metal, rather than some variant of net lifetime earnings.

[19] Golf, the other major lucrative individual sport, lacks global junior ranking data. Chess offers better measures of individual quality, but lacks sufficient compensation to attract potential grandmasters from high income countries (see "Stairway to Heaven").

[20] The International Tennis Federation began ranking age 18 and under (U18) players in 1978. The Tennis Europe Junior Tour has published age 14 and under (U14) and age 16 and under (U16) year-end rankings since 1990, which includes players from around the world, although most are European.

[21] Due to the difficulty of measuring individual abilities in team sports, youth ability measures typically constitute the subjective judgements of scouts. Some limited scouting data of youths exist, but have not yet been used to predict professional outcomes. For basketball, two scouting agencies rank high school basketball players, including freshmen to seniors but with many more older students, based on perceived quality from 2002-2012 (see http://www.rivals.com) associated with Yahoo! Sports and http://www.scout.com associated with http://www.foxsports.com (see Anderson and Sinkey 2013). Although baseball has long included high school students in scouting reports and have drafted high school graduates (aged 17+), our analysis provides indicators of professional success at much earlier ages.

For players as young as age 13, who have received objective signals about their relative international performances and, hence, potential to earn a living on the professional tennis tour, we estimate players' lifetime earnings distributions from their annual global rankings. Based on those estimated career earnings, we analyze players' decisions to continue making incremental human capital investments each year or to quit and focus on other pursuits. The winner-take-all nature of professional tennis gives rise to extremely skewed earnings distributions and we test whether this skewness influences players' participation decisions, along with the mean and variance of lifetime earnings.

We find that although teenage players prefer high mean earnings and a low variance, they are also attracted to highly-skewed earnings distributions – like gamblers at horse races or in lotteries. If skewness were to fall to zero, boys would be 23% less likely on average to continue in the sport, whereas girls would be 5% less likely to continue. A side-effect of winner-take-all labor markets therefore is an increase in participation, even among those who have a negligible chance of earning large amounts.[22]

We offer two other major contributions in addition to creating the first youth-to-professional domain data set of abilities and earnings and incorporating skewness preferences into occupational choice decision-making. The first of those is that we compare our analysis of high stakes career gambles and of low stakes lotteries and find 20 times greater skewness preferences for low stakes lottery players than for high stakes career lotto gambles. Finally, our results reveal that boys exhibit four times the skewness preference of girls in these exceedingly risky occupational choices.

The paper proceeds as follows. The next section provides background information about decision-making under uncertainty. Then we develop a simple model that incorporates the role of skewness in career choice. In section 4 we introduce our and illustrate the relationships between lifetime earnings and pre-professional youth rankings. Section 5 presents our main empirical results and in section 6 we compare the role of skewness attraction in our high stake career choice decisions compared to small stake lottery gambles, derived from Rieger et al. And then we conclude.

[22] For an anecdotal account, see "Thousands of Players, Hundreds of Events and Little Reward," (David Waldstein, *New York Times*, August 18, 2017).

2. Background

Two broad sets of theories have been proposed to explain decision-making under uncertainty. In a sketch of a standard rational expectations expected utility model, Rosen and Sanderson (2001) characterize winner-take-all labor market participants as dynamic learners about their abilities and prospects of success, by regularly reassessing the expected value of their lifetime earnings or their chances of elite employment, for example by a major-city orchestra or a professional sports team.

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 the superstar model by MacDonald (1988) 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 the superstar model of Rosen (1986) and the occupational choice model of Miller (1984). 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".

Although expected utility theory assumes that all percentage points of risk are equally important, prospect theory proposes that the values of outcomes of risky prospects are 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 1979, 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).

Although entrants to superstar markets are often assumed to be risk lovers, the considerable empirical literature on gambling on horse races and lotteries finds a willingness to accept a lower expected payoff in return for greater skewness, rather than being risk loving (Golec and Tamarkin 1998, Clotfelter and Cook 1993, Garrett and Sobel 1999). The analogy with lotto is useful in thinking about the distribution of incomes as a matrix of payoffs. Clotfelter and Cook (1993) find lotto sales to be positively related to the size of the jackpot, but negatively to expected value. Golec and

Tamarkin (1998) find that horse track bettors are attracted to the positive skewness of returns offered by low probability, high variance bets, rather than being risk lovers with mean-variance utility functions, as suggested by Quandt (1986). Garrett and Sobel (1999, p.88) conclude "that lottery players, like horse race bettors, are risk averse but favor positive skewness".[23]

Although most of the empirical and experimental studies cannot distinguish between preferences for positive skewness versus miscalculations about the probability of such outcomes, Snowberg and Wolfers (2010) attempt to distinguish between preference-based and perception-based explanations and find evidence that misperceptions of probabilities account for the longshot bias among horse track gamblers.

Whereas lotto and lottery games entail purely random outcomes, we analyze winner-take-all labor market results based on individual abilities. What has remained elusive for understanding labor supply in winner-take-all markets are longitudinal datasets that link youth and adult relative performances to career earnings.

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), if in college, whether to continue or to dropout (Stange 2012, Stinebrickner and Stinebrickner 2012), and the choice of a college major (Arcidiacono *et al*. 2012). Generally, these authors find a reasonable approximation to Bayesian updating based on sequential experiences that provide new information about individuals' match with particular training programs.

3. Model

Superstar markets are binary since, for example, drafted players and other superstars earn in the NBA and non-drafted players earn nothing and typically do not have ancillary careers, such as coaching, that offer lifetime incomes comparable to that of a college educated worker. Although an elite group of "superstars", for example, accountants, lawyers or dentists who earn standard deviations greater than others, those professions do no constitute winner-take-most markets in our view, as Frank and Cook

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

(2010) suggest, because the typical professional in those fields earns a positive rate of return on their human capital investment.

Assume that people have a choice between a risky career, in which their future earnings are uncertain, and a riskless career, which pays a given amount with certainty. People work as apprentices for *T* periods before their final earnings, *w*, are revealed. Wages during the apprenticeship phase are assumed to be the same on the risky and riskless job.

Final earnings each period on the risky job are determined by a person's ordinal ranking according to performance, which is not known in advance. However, at the end of each period *t*, people learn their current ranking, *r*.

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

$E_t(u(w_i)) > u(\widehat{w}_i)$,

where *u* is the person's utility function over lifetime income, *w* is lifetime income on the risky career, $\widehat{w}$ is lifetime income on the riskless career.

The person's utility function can be approximated by a Taylor series expansion around the mean of *w*, $\overline{w}$, (Golec and Tamarkin 1998):

$$u(w_i) \approx u(\overline{w}) + u'(\overline{w})(w_i - \overline{w}) + \frac{u''(\overline{w})}{2}(w_i - \overline{w})^2 + \frac{u'''(\overline{w})}{6}(w_i - \overline{w})^3 . \qquad (1)$$

Then, taking expectations at time *t* and adding an error term:

$$E_t(u(w_i)) = u(\overline{w}) + u'(\overline{w})E_t(w_i - \overline{w}) + \frac{u''(\overline{w})}{2}E_t(w_i - \overline{w})^2 + \frac{u'''(\overline{w})}{6}E_t(w_i - \overline{w})^3 + \varepsilon_{it} .$$

Let $\varepsilon_{it} \sim U[-e,e]$. Then the probability that the person will continue on the risky career after period *t* is:

$$\begin{aligned} P(\text{risky career}_{it}) &= P(\varepsilon_{it} > u(\hat{w}) - u(\overline{w}) - u'(\overline{w})E_t(w_i - \overline{w}) - \frac{u''(\overline{w})}{2}E_t(w_i - \overline{w})^2 \\ &\qquad - \frac{u'''(\overline{w})}{6}E_t(w_i - \overline{w})^3) \\ &= \frac{e - u(\hat{w}) + u(\overline{w})}{2e} + \frac{u'(\overline{w})}{2e}E_t(w_i - \overline{w}) + \frac{u''(\overline{w})}{4e}E_t(w_i - \overline{w})^2 \\ &\qquad + \frac{u'''(\overline{w})}{12e}E_t(w_i - \overline{w})^3 . \qquad (2) \end{aligned}$$

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 can 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.

In each period of the apprenticeship phase, a person's expectations depend on his/her rankings up to that point. Therefore, equation 1 can be rewritten as follows:

$$P(\text{risky career}_{it}) = \frac{e-u(\hat{w})+u(\overline{w})}{2e} + \frac{u'(\overline{w})}{2e}E(w_i-\overline{w}\mid \mathbf{r}_{it}) + \frac{u''(\overline{w})}{4e}E((w_i-\overline{w})^2\mid \mathbf{r}_{it}) + \frac{u'''(\overline{w})}{12e}E((w_i-\overline{w})^3\mid \mathbf{r}_{it}), \qquad (3)$$

where $\mathbf{r}_{it}$ is the sequence of person i's rankings up to period t.

4. Data

Our dataset combines four distinct sources for tennis rankings and earnings, all of which are available for females and males. Overall, we include every player born between 1977 and 1986 who appeared at least once in an international ranking between the ages of 13 and 30. First, since its inception in 1990, the Tennis Europe Junior Tour publishes year-end rankings for U14 and U16 players from around the globe who compete in numerous tournaments throughout Europe. These data provide each player's full name, country of origin, birthday, and ranking. This Tour represents the earliest and most comprehensive rankings available for tennis players competing in international events.

Second, we access the U18 worldwide rankings, published by the International Tennis Federation (ITF). For a young potential entrant into professional tennis, this Tour provides the next and final step before entering the professional arena. Similar to Tennis Europe, the ITF data include each player's full name, birth date, nationality, and ranking. Third and final, data regarding players' professional performances comes from the respective professional organizations: the Women's Tennis Association (WTA) and the Association of Tennis Professionals for men (ATP). To derive a comparable universe of players throughout all age and ranking categories over time, our analysis focuses on the cohort of players born between 1977 and 1986. This timeframe guarantees observing players' performances from age 13 (when Tennis Europe rankings became available) to 30 (since we record players' professional performance until the end of 2016). Our dataset includes any player born in that age cohort who shows up at least once in any of the above categories of rankings.

In his review of *The Winner-Take-All Society* by Frank and Cook (1995), Rosen (1996) argues that "few seriously try to enter these professions . . . and those that do are predominantly unskilled and troubled people who never played the game at all" so that the "the inefficiencies they [Frank and Cook] claim seem to me be greatly exaggerated" (134).

Figure 1 plots a player's annual prize money in the WTA or ATP against his/her year-end ranking. This illustrates how closely related players' earnings are to their ranking, as well as the winner-take-all nature of professional tennis. Because prize money is allocated according to a player's position in each tournament and because this declines sharply with tournament position, the top male and female players account for the vast majority of total prize money.

The decision facing young players is 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 the expected earnings. However, while lifetime prize money is highly non-linearly related to junior ranking (as seen in the bottom panel of Figure 2), a person's lifetime number of tournaments is roughly linearly related to ranking (as seen in the top panel).

Our analysis requires estimates of the lifetime prize money distribution a given player expects to face. In each year, we calculate a player's ranking within the cohort of players of the same sex born in the same year, *b*, in either the U14, U16 or U18 age categories or the professional tour.[24] A player's total future prize money, *W*, is determined by his/her age-specific ranking in each future year, *r*:

$$W_{it} = W(\mathbf{r}_i) = \sum_{v=1}^{b+30} w(r_{i(t+v)}) \,. \tag{4}$$

We assume that players only receive prize money between ages 19 and 30. Since players do not know the future prize money distribution, we assume that they use the

[24] 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.

distribution of prize money on the professional tour (*i.e.* WTA and ATP) in the previous year as a reference point:

$$w(r_{i(t+v)}) = w(r_{t+v}, a_{it} + v, s_i, t-1), \tag{5}$$

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. To simplify the computation, 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:

$$P(W_{it}) = P(r_{i(t+30-a)}, r_{i(t+29-a)}, \ldots, r_{i(t+1)}). \tag{6}$$

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:

$$P(W_{iit}) = P(r_{t+v} \mid r_{t+v-1}, a_{it} + v - 1, s_i) P(r_{t+v-1} \mid r_{t+v-2}, a_{it} + v - 2, s_i) \ldots P(r_{t+1} \mid r_{it}, a_{it}, s_i). \tag{7}$$

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. To simplify the computation, we group players into 10-ranking bands between rankings 50 and 500 and 100-ranking bands above ranking 500.

Combining the age-specific transition probabilities and the age-specific earnings distributions and considering every possible rank in every future year of a player's career, we can calculate lifetime prize money PDFs for each player in each year. These exhibit significant variation. Figure 3 plots the distributions faced by 18-year olds in 1997 (near the middle of our sample) with different rankings. Even those top- ranked in their age group face a reasonably high chance of earning very little over their careers, especially boys. However, those ranked 100 experience a much higher likelihood that players will earn close to zero over their careers. This fact is reflected in the skewness coefficients for distributions (calculated as $E(W_i - \overline{W})^3 \big/ (E(W_i - \overline{W})^2)^{3/2}$ and reported under each histogram), which are much higher for those ranked 100 than for those ranked 1.

As players age, their rankings become better and 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 have a very high chance of making little money over their careers, compared to top-ranked 18-year-olds. Accordingly, the skewness coefficient falls with age.

We can calculate three moments of the lifetime earnings distributions for any player *i* in any year *t*:

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

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

$$E(W_{it} - E(W_{it}))^3 = \sum_{W} (W_{it} - E(W_{it}))^3 \quad . \tag{8}$$

A player is considered to be active in tennis in a given year if he/she participated in any of the tennis divisions in our dataset (U14, U16, U18 or the professional tours). The means for the primary variables of interest are reported in Table 1. Mean expected prize money is slightly higher for girls than for boys, even though professional men are paid more than equal-ranked women, because more boys participate in junior tennis and they face a higher chance of earning very little over their careers. Boys also have a much greater expected variance and skewness than girls.

5. Estimates of the model of tennis participation

A dummy variable for whether a player was active is regressed on the mean, variance and skewness of his/her expected career prize money, given his/her ranking in the previous year, consistent with equation 3, as follows:

$$P(\text{active}_{it}) = \alpha_1 E_{t-1}(W_i) + \alpha_2 E_{t-1}(W_i - \overline{W}))^2 \\ + \alpha_3 E_{t-1}(W_i - \overline{W}))^3 + \mathbf{AGE}_{it}\boldsymbol{\beta} + \gamma_i + \varepsilon_{it}\,, \tag{9}$$

where **AGE** is a full set of age dummies, intended to capture the effects of changes in opportunity cost over a person's teenage years, γ is a player fixed effect (which is omitted in some specifications) and ε is a random error term.

The results for boys are reported in Panel A of Table 2. As predicted by theory, the estimates of α_1, α_2 and α_3 are positive, negative and positive, respectively. In the first column, the 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 boy would be 23 percent less likely to continue playing tennis the following year. The age coefficients indicate that players are increasingly likely to quit tennis as they age, with an especially large fall in participation at age 19, when most players have to decide whether to attend college.

The results for girls are reported in Panel B of Table 2. The coefficients in the first column imply an elasticity with respect to the mean of 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, meaning that girls would be 5 percent less likely to stay in tennis on average if skewness fell to zero.

The results in the first column of Table 2 do not take into account the cost of competing in professional tennis. Since top-ranked players travel more widely and require more support staff than lower-ranked players, our estimates may overstate the skewness in lifetime earnings. To address this, we subtracted an estimate of cost-adjusted prize money, provided to aspiring players by the International Tennis Federation. This varies by sex, continent and ranking.[25] As seen in the second column of Table 2, adjusting for costs makes little difference to the results. The skewness elasticities fall by the same magnitude for boys and girls, but loses significance for girls given the small amount of skewness affection exhibited by 14-19 year olds (Table 2, panel B, column 1).

In the third column of Table 2, we repeat the specification in the first column, using observations for ages 20-28. These players are already in the professional tour and are considering whether to continue or quit, based on their current ranking. We find much smaller skewness effects and the coefficient for women is now insignificant. Over the full 14-28 age range, however, both men and women exhibit skewness loving behavior (as seen in the last column of Table 2).

6. Heterogeneity in skewness preferences by size of gamble

[25] Since we have no information on a player's current place of residence, we use their continent of birth. For 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.

Whereas lotto and lottery games entail small stakes that are purely random outcomes, we analyze winner-take-all labor market results based on individual abilities. How might one expect skewness preferences to compare for lotteries and career lotto? Although a person might purchase a negative expected return lottery ticket for \$2 for the entertainment value of the fantasy of hitting the jackpot, career lotto choices may alter or fundamentally limit a person's occupational opportunity set. Nonetheless, despite the stakes being much higher, the results discussed above suggest that individuals approach the choice of career in a manner reminiscent of how they approach lotteries.

To examine how our estimates of the magnitude of the skewness effect compare with those exhibited in smaller lotteries, we compare our results with those obtained from the experimental data collected by Rieger *et al*. (2015). Rieger *et al*. asked participants in laboratory experiments in college classrooms in 52 countries to give certainty equivalents for a series of hypothetical lotteries in which they stood to gain at most \$10,000 or to lose at most \$100, with an average payoff of around \$800. Although the authors did not examine attitudes towards skewness, estimates of the relationship between the certainty equivalent and the skewness can nonetheless be derived from their data.[26]

Each lottery has 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 payouts, so that the payouts have mean zero and variance one in each case.

In general, person *i*'s risk premium for lottery *j* can be written as follows:

$$\overline{w}_j - y_{ij} = \frac{u''(\overline{w}_j)}{2u'(\overline{w}_j)} E(w_j - \overline{w}_j)^2 + \frac{u'''(\overline{w}_j)}{6u'(\overline{w}_j)} E(w_j - \overline{w}_j)^3, \tag{10}$$

where y is the person's certainty equivalent, w is a possible payoff for the lottery and $\overline{w}$ is the mean payoff. Therefore, the risk premium for standardized lottery *j* can be written:

$$\frac{\overline{w}_j - y_{ij}}{\sqrt{E(w_j - \overline{w}_j)^2}} = -\frac{u''(0)}{2u'(0)} - \frac{u'''(0)}{6u'(0)} E\left(\frac{w_j - \overline{w}_j}{\sqrt{E(w_j - \overline{w}_j)^2}}\right)^3. \tag{11}$$

[26] These are available at http://dx.doi.org/10.1287/mnsc.2013.1869.

Equation 11 implies that the standardized risk premia are linearly related to the skewness of a lottery. Therefore, we estimated this relationship using OLS, allowing for person fixed effects. For comparability with our tennis estimates, we used only those aged 16-19 (although this made little difference to the results).[27] The slope coefficients provide estimates of a person's willingness to pay for a unit of skewness, or skewness preference (which will be negative if people are skewness loving), and the intercept provides an estimate of a person's risk preference, specifically the Arrow-Pratt measure of risk aversion divided by two (which will be positive if people are risk averse).

Comparing equation 3 with equation 11, it is clear that the ratio of the coefficient on the skewness and 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 preference for skewness, *relative* to their preference for the mean (that is, 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*.), 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.

As reported in Table 3, the Rieger *et al*. data reveal a relative skewness preference for boys of 0.441 and a relative skewness preference for girls of 0.473, which are 21 and 14 times larger than the corresponding estimates from the tennis career decision, respectively.[28] Hence, skewness preferences have much larger effects 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 less influenced by long-shot outcomes (even though a preference for low-probability, high-return outcomes still has a sizeable effect on behavior, as seen in Table 2).

For both datasets, we also combined observations for both genders and allowed interactions with age and the GDP quintile of a person's country of birth. Relative skewness preference declines slightly with age in both datasets. Although there is a

[27] Rieger *et al*. did not include anyone under 16 in their study.

[28] Note that although we found in Table 2 that the overall skewness elasticity was smaller for girls than for boys, since girls were also much less sensitive to the mean, their *relative* skewness preference is larger than boys', as also found in the Rieger *et al*. data.

monotonic relationship between GDP and relative skewness preference over modest-sized lotteries, we find a U-shaped relationship in the career choice setting. Young players from the poorest countries are relatively attracted to highly-skewed earnings distributions, unlike what we found with the Rieger *et al*. data.

7. Conclusion

The fundamental dilemma with labor supply decisions in winner-take-all markets is that the beauty, fame and fortune of superstars entices potential entrants to devote substantial upfront skill-specific human capital investments without data to allow performers, parents and coaches to calculate meaningful success probabilities. The public policy concern regarding aspiring participants in these markets is the significant individual and social inefficiency that results from the resources spent developing skills for low probability, high payoff careers, rather than investing in more realistic occupational paths. Ample behavioral economics analyses indicate the difficulty young agents (aspiring performers) and emotionally-involved ones (*i.e.* parents) have functioning as rational Bayesian updaters when making uncertain intertemporal career decisions.

Here we address the puzzle of why so many people play career lotto by making large financial occupation choice gambles in exceedingly risky, long-shot superstar markets. Such labor markets are characterized by extremely skewed earnings distributions. We offer the first empirical analysis of labor supply decisions under risk in winner-take-all markets by creating a unique longitudinal dataset with objective global rankings of youth and of professional participants and of career earnings. Our data permit estimation of the rational expectations expected utility model suggested by Rosen and Sanderson (2001): do winner-take-all labor market participants act as dynamic learners about their abilities and prospects of success, regularly reassessing the expected value of their lifetime earnings or their chances of elite employment, for example by a major-city orchestra or a professional sports team, and exit when the prospects become too low?

Using longitudinal data on aspiring and actual professional tennis players, we estimate a lifetime prize money distribution for each player, given his/her age, current ranking and the observed distribution of prize money among current professionals. We find that teenagers are more likely to stay in tennis if they face a distribution with a high mean, low variance and high skewness. All else equal, boys would be 23% less and girls 5% less likely to continue in tennis each year if the skewness of each player's lifetime prize money distribution were reduced to zero. The fact that players are attracted to highly-skewed distributions is reminiscent of the behavior of gamblers at horse races or of lottery entrants. Comparing our results with experimental data on modest-sized lotteries, we find that people exhibit a much smaller preference for skewness when choosing careers. Nonetheless, the magnitude of our estimates suggests that a love of skewness, among boys especially, can significantly increase labor supply in winner-take-all markets.

References

Anderson, B. C., and M. J. Sinkey (2013). *Like Mike or like LeBron: Do the most able need college to signal?* Working paper.

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

Baker, J., S. Cobley, J. Schorer, and N. Wattie (2017). *Routledge Handbook of Talent Identification and Development in Sport*. Abingdon: Routledge.

Clotfelter, C. T., and P. J. Cook (1993). The peculiar scale economies of lotto. *American Economic Review* 83 (3), 634-643.

Dertwinkel-Kalt, M., and M. Koster (2017). Salient compromises in the newsvendor game. *Journal of Economic Behavior and Organization*, forthcoming.

Ericsson, K. A., N. Charness, P. J. Feltovich, and R. R. Hoffman, eds (2006). The Cambridge Handbook of Expertise and Expert Performance. Cambridge: Cambridge University Press.

Frank, R. H., and P. J. Cook (1995). *The winner-take-all society: Why the few at the top get so much more than the rest of us*. New York: Penguin.

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

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

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

Kahneman, D., and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2), 263-292.

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

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.

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

Lubinski, D., C. P. Benbow, and H. J. Kell (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), 155-166.

Marshall, A. (2009). *Principles of economics: Unabridged eighth edition*. New York: Cosimo Classics.

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

NPR (2015). *Sports and health in America*. http://www.rwjf.org/content/dam/farm/reports/reports/2015/rwjf420908. Accessed: 2015-09-21.

Quandt, R. E. (1986). Betting and equilibrium. *Quarterly Journal of Economics* 101 (1), 201-207.

Rieger, M. O., M. Wang, and T. Hens. 2015. Risk preferences around the world. *Management Science* 61 (3), 637-648.

Rosen, S. (1981). The economics of superstars. *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 A. Sanderson (2001). Labour markets in professional sports. *Economic Journal* 111 (469), 47-68.

Savage, L. J. (1972). *The foundations of statistics*. New York: Dover Publications.

Schonbrun, S. (2017). "Shalane Flanagan Solves N.Y.C. Marathon for American Women," *New York Times*, November 5, F1.

Slovic, P., and A. Tversky (1974). Who accepts Savage's axiom? *Behavioral Science* 19 (6), 368-373.

Smith, A. (1937). *An Inquiry into the Nature and Causes of the Wealth of Nations*. New York: Modern Library.

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

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

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

Terviö, M. (2009). Superstars and mediocrities: Market failure in the discovery of talent. *Review of Economic Studies* 72 (2), 829-850.

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

Figure 1

Year-end professional ranking and prize money in 1997

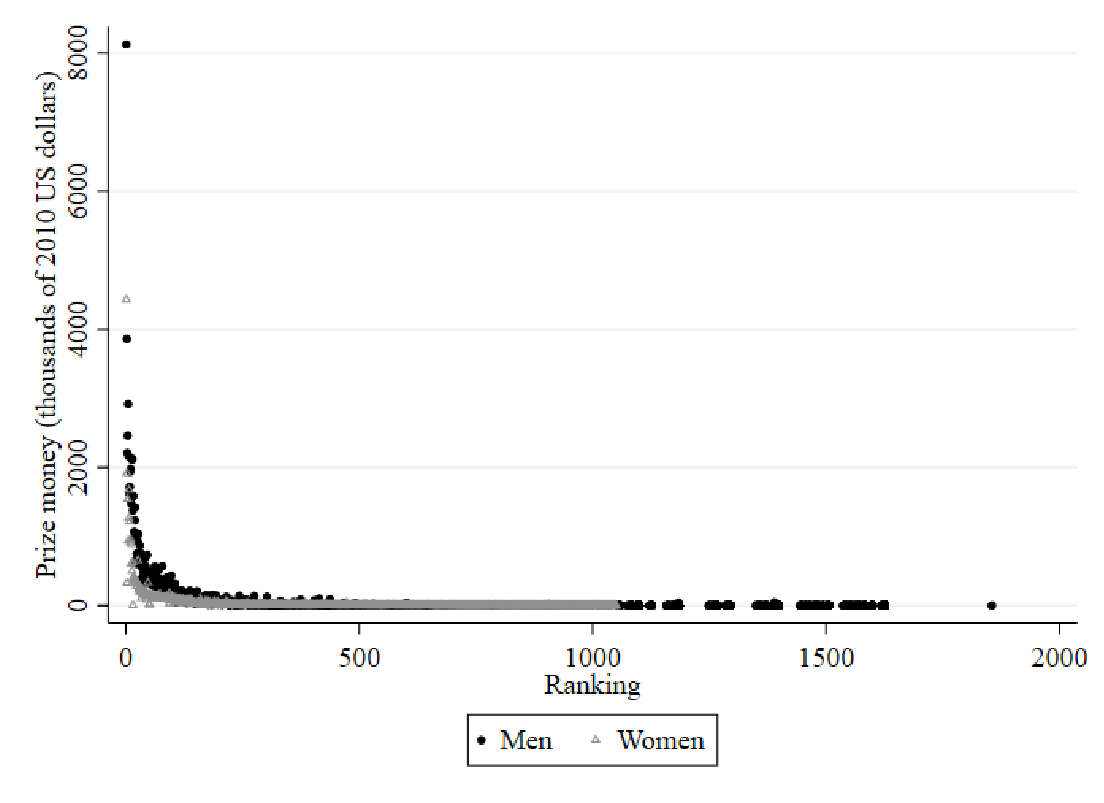


Figure 2

Lifetime number of tournaments and average prize money by ranking at age 18

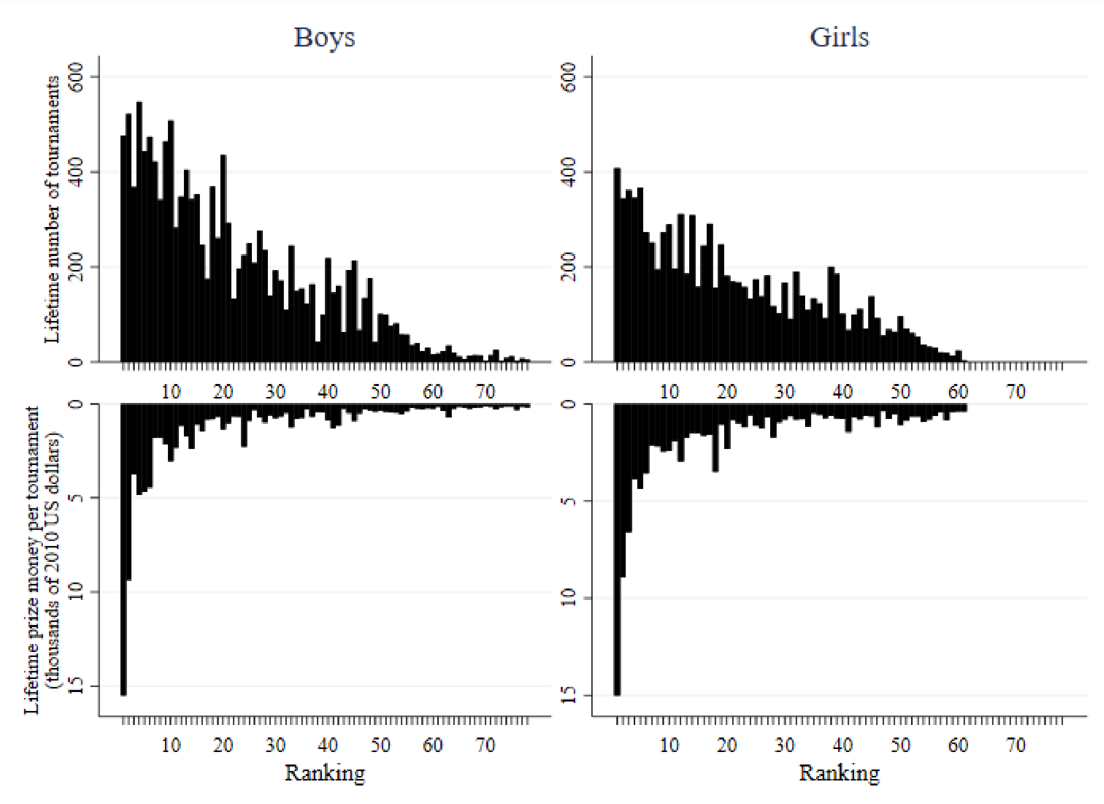
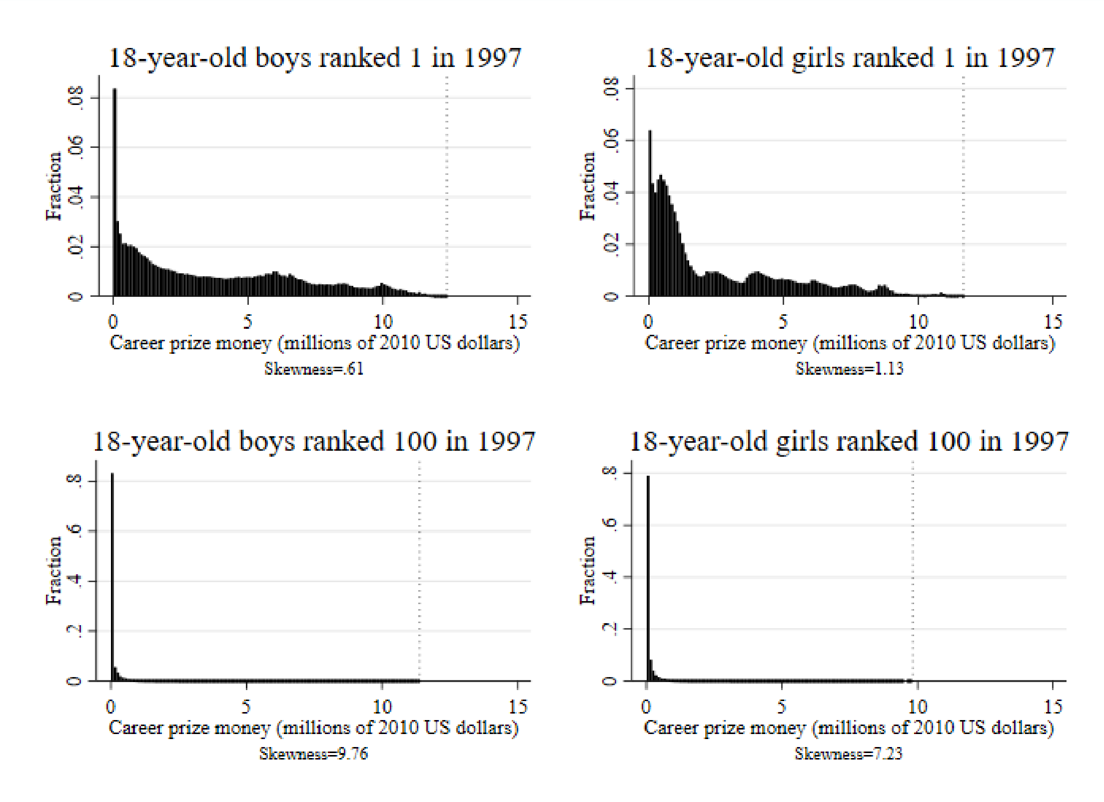


Figure 3

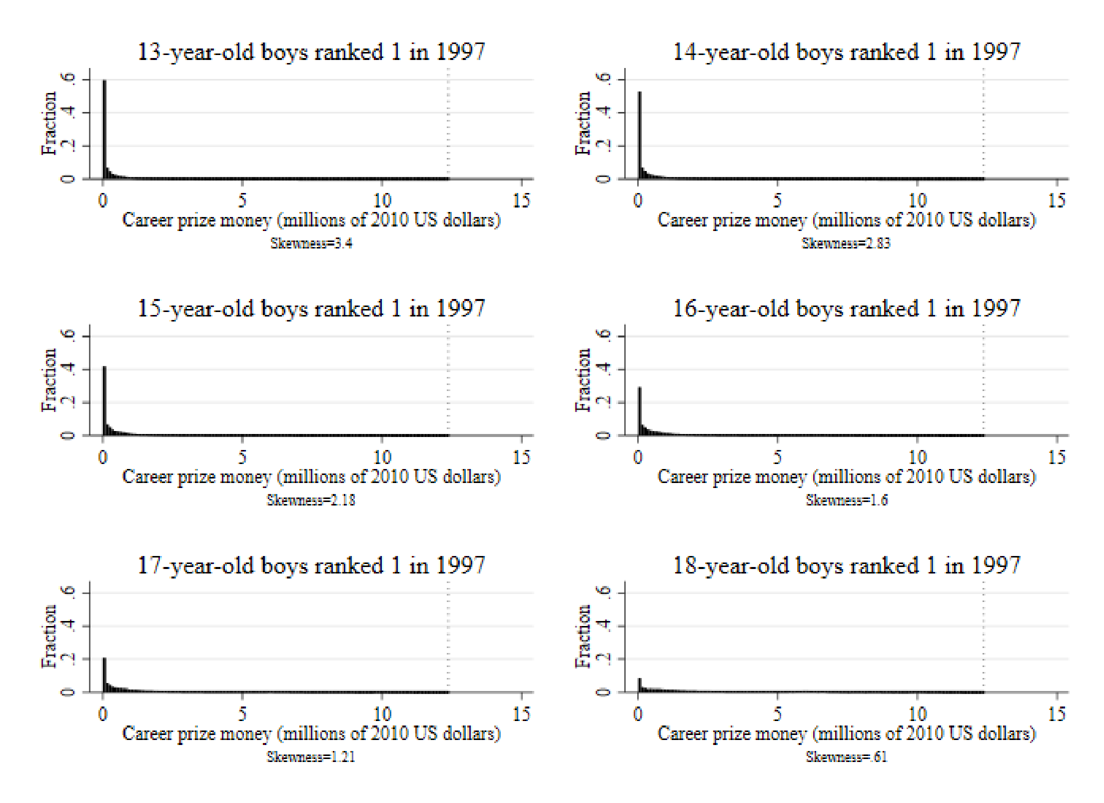
Histograms of predicted career prize money for 18-year-olds in 1997



Note: The dotted line denotes the maximum career prize money.

Figure 4

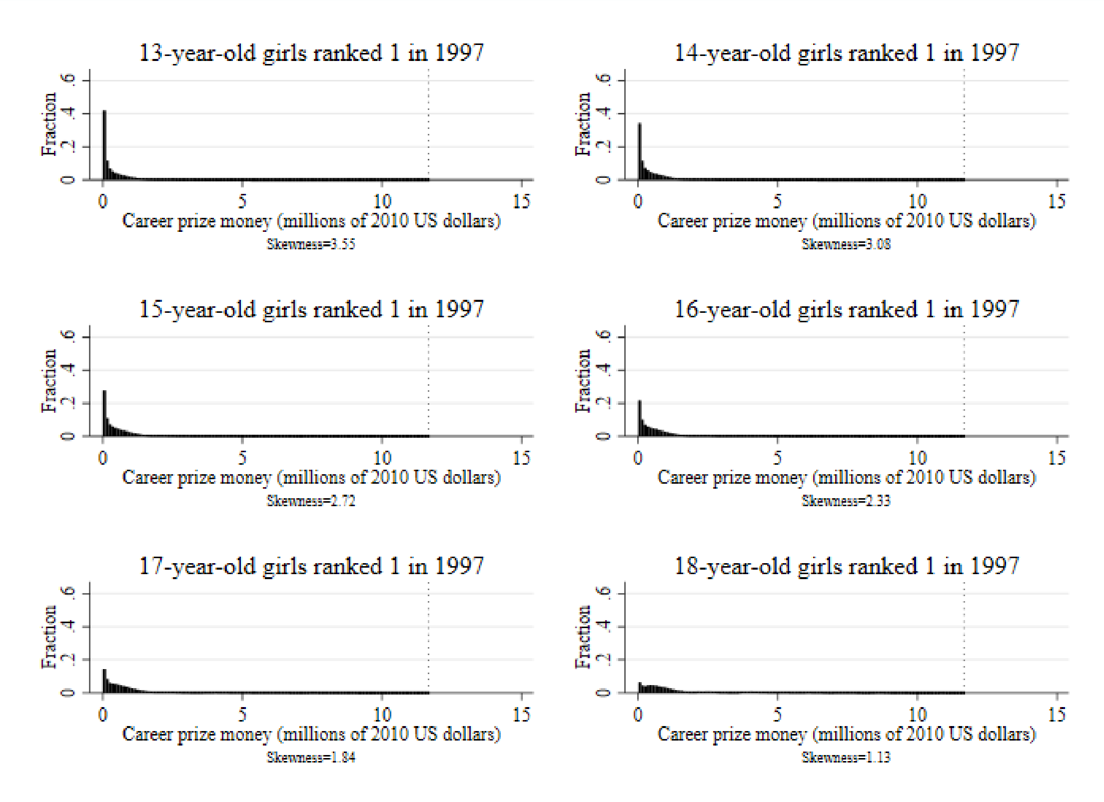
Histograms of predicted career prize money for boys in 1997



Note: The dotted line denotes the maximum career prize money.

Figure 5

Histograms of predicted career prize money for girls in 1997



Note: The dotted line denotes the maximum career prize money.

Table 1

Means of the key variables

Variable	Boys	Girls
Active	0.552	0.657
Lagged mean of predicted career prize money	0.290	0.298
Lagged variance of predicted career prize money	0.899	0.499
Lagged skewness of predicted career prize money	5.653	2.143
Lagged ranking	119.150	87.029
Age	16.862	16.679
Number of observations	15,810	15,078

Note: The mean of career prize money is expressed in millions, the variance in trillions and the skewness is in quintillions of 2010 US dollars.

Table 2

Results of estimating participation equation

A. Boys

Variable	Ages 14-19		Ages 20-28	Ages 14-28
	(i)	(ii)	(iii)	(iv)
Lagged mean of career prize money	1.162***	0.904***	0.119***	0.160***
(millions of 2010 US dollars)	[0.610]	[0.179]	[0.062]	[0.083]
	(0.082)	(0.097)	(0.026)	(0.022)
Lagged variance of career prize money	-0.341***	-0.261***	-0.053***	-0.055***
(trillions of 2010 US dollars)	[-0.554]	[-0.386]	[-0.050]	[-0.076]
	(0.031)	(0.033)	(0.013)	(0.011)
Lagged skewness of career prize	0.022***	0.019***	0.004**	0.008***
money (quintillions of 2010 US	[0.228]	[0.179]	[0.014]	[0.063]
dollars)	(0.002)	(0.002)	(0.002)	(0.001)
R-squared	0.618	0.615	0.598	0.548
Number of observations	15,810	15,810	6,009	21,819

B. Girls

Variable	Ages 14-19		Ages 20-28	Ages 14-28
	(i)	(ii)	(iii)	(iv)
Lagged mean of career prize money	0.608***	0.237***	0.246***	0.372***
(millions of 2010 US dollars)	[0.276]	[-0.017]	[0.098]	[0.163]
	(0.066)	(0.092)	(0.043)	(0.031)
Lagged variance of career prize money	-0.192***	-0.014	-0.104***	-0.130***
(trillions of 2010 US dollars)	[-0.146]	[-0.009]	[-0.046]	[-0.087]
	(0.033)	(0.048)	(0.023)	(0.017)
Lagged skewness of career prize	0.014***	-0.000	0.001	0.012***
money (quintillions of 2010 US	[0.045]	[-0.001]	[0.001]	[0.033]
dollars)	(0.002)	(0.006)	(0.004)	(0.002)
R-squared	0.643	0.641	0.558	0.560
Number of observations	15,078	15,078	4,693	19,771

Notes: 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%, 5% and 1% level, respectively.

Table 3

Relative skewness preferences by demographic group

Population group	Gamble	
	Continue tennis	Rieger *et al*. lotteries
Men	0.019***	0.404***
Women	0.033***	0.473***
Age 16	0.055***	0.465***
Age 17	-0.011	0.468***
Age 18	0.027***	0.441***
Age 19	0.024***	0.440***
Birth country GDP quintile 1	0.031***	0.165***
Birth country GDP quintile 2	0.029***	0.273***
Birth country GDP quintile 3	0.013*	0.401***
Birth country GDP quintile 4	0.018***	0.472***
Birth country GDP quintile 5	0.025***	0.438***

Notes: Values in the first column are 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 are coefficients on the skewness of a lottery from a regression of equation 11.

*, ** and *** denote significance at the 10%, 5% and 1% level, respectively.