Like Mike or Like LeBron: Do the Most Able Need College to Signal?

Benjamin Anderson [*], Michael Sinkey [†]

December 22, 2016

Abstract

Do individuals with demonstrably high ability need to attend college to further signal their ability to potential employers? We examine the labor market entry decision for basketball players deciding to enter or return to college versus entering the labor market for professional basketball, specifically the National Basketball Association (NBA). Individuals in this market have significant financial incentive to forgo further schooling in order to pursue their careers immediately and therefore face a trade-off between possible immediate financial rewards and the acquisition of additional skill-related human capital or improving the signals regarding own productivity. We exploit variation in the strength of signals regarding own ability that players receive while in high school from external ratings agencies, Rivals and Scout. Players deemed to be of the highest ability in each high school graduating cohort receive both a continuous, ordinal ranking as well as a categorical, "star" rating. After providing descriptive evidence for a regression probability jump kink design (RPJK), we estimate the probability of collegiate prospects entering into the professional labor market as a function of the strength of ability signal that players receive while in high school. In our preferred specifications, a one-unit increase in ranking for players receiving the highest "star" ranking is associated with a 3.8 percentage point increase in the probability of entry for college freshman relative to comparable players at the threshold that receive the next highest ranking, which is equivalent to an increase in one-fourth of a standard deviation in the probability of entry. Our results suggest that, more generally, only a very small fraction of the most able labor market participants would forgo schooling based on the intensity of a signal of ability.

[*]Department of Economics, Colgate University, e-mail: bcanderson@colgate.edu

[†]Department of Economics, University of West Georgia, 1601 Maple Street, Carrollton, GA 30118, e-mail: msinkey@westga.edu

1 Introduction

The classic labor market signaling model formulated by Spence (1973) [26] posits that high ability individuals may pursue secondary education regardless of whether any human capital accumulation takes place in order to signal high productivity to employers. Empirically, it is well established that individuals may receive higher wages from credentials that may only be loosely related to human capital accumulation and may solely wish to communicate high ability and future productivity (see Hungerford and Solon (1987) [17]; Weiss (1995) [30]; Jaegar and Page (1996) [18]; Tyler, Murnane, and Willett (2000)[29]; Flores-Lagunes and Light (2010) [11], among others). However, a majority of labor market entrants do not possess publicly-observable signals of productivity before entering college. Once the decision to attend college has been made, it is difficult to separately identify gains from productivity-enhancing human capital and gains from additional signals of productivity. This paper explores the decisions of individuals with observable signals of high productivity to pursue additional years of schooling relative to entering the labor market.

We examine the schooling and labor market entry decisions of workers in a particular setting–professional basketball–when publicly-available signals of productivity are generated in high school. Players that wish to enter the basketball labor market must first enter a draft, where professional teams take turns selecting new workers based on expected productivity. If a player believes that a professional team will find him to be sufficiently productive to be selected, then he has little fiduciary incentive to pursue more schooling, since the expected value of the initial contract exceeds the option value from remaining in school. Using variation in the quantity and quality of productivity signals across high school players, we estimate the probability of entry into the draft, the likelihood of being selected in the draft, and the value of that selection to the player. We then exploit an exogenous discontinuity in signal strength to determine the impact of information about expected worker productivity on labor market outcomes.

It is well-known that the study of labor market outcomes can be confounded by the presence of unobservable worker characteristics, such as soft skills, non-cognitive skills, personality traits, and external connections, which can jointly influence workers' entry decisions, firms' hiring decisions, and workers' long-term wages (Bowles, Gintis, and Osborne, (2001) [7]; Nyhus and Pons, (2005)

[25]; Borghans, ter Weel, and Weinberg, (2008)[6]. Although some of the skills relevant to the professional basketball labor market remain unobservable, the majority of the most important factors influencing both worker and firm actions are observable. We observe a wide variety of performance metrics for college basketball players directly correlated with player performance in the NBA. Coupling these metrics with external assessments of ability allows us to plausibly control for most relevant determinants of a player's selection. Moreover, the collectively-bargained provisions governing both the length and salary structure of worker contracts and worker movements between teams enables us to distinguish between alternative explanations for schooling choices, allowing the decision to enter the NBA draft to be plausibly framed as an option value decision for enrollment, as in Stange (2012) [27]. Finally, the economic importance of superstars in the NBA, identified by Hausman and Leonard (1997) [15], creates incentives to identify these individuals at an early age for teams, advertisers, as well as the players themselves.

For identification, we concentrate on the labor market entry decisions of players whose observable productivity characteristics are exchangeable, but whose external signals of ability vary discontinuously. Utilizing a regression probability jump kink (RPJK) design, we estimate the probability of collegiate prospects entering into the professional labor market as a function of the strength of ability signal that players receive while in high school. In particular, we utilize the variation in signal strength regarding own ability from two different external ratings agencies–Scout, and Rivals. Identification comes from the fact that both agencies offer two sets of rankings for each player. The first is a granular, ordinal ranking in which the top 100 (Scout) and top 150 (Rivals) players are put in order based on perceptions of ability. The second set is an aggregated ordinal ranking that groups players by "star" into "three-, four-, and five-star" categories, which roughly correspond to a "good, better, best" assessment of ability. We examine the entry decisions of players with adjacent granular rankings, but distinct aggregated rankings. Aside from a different aggregate assessment of ability, these players are exchangeable.

We find that, conditional on being assigned into the highest ability group, players use the information contained in the granular signal to enter the draft far more aggressively. Because one is the best possible ranking a player can have, we find that a one-slot reduction in ranking (the player is closer to being considered the best) leads to between a 1.6 and 3.8 percent increase in the probability of entering the draft, which is equivalent to one-fourth of a standard deviation of the probability of entry within our sample. Players that are assigned into the *next* group do not exhibit any such sensitivity to their granular ranking, and, in fact, enter the draft very infrequently.

We find that our results are more pronounced for *new* potential labor market entrants. In our sample, the highest kink estimates are for freshmen, who possess relatively little information about own productivity relative to players who have played multiple seasons in college. We interpret these findings as being suggestive of the notion that initial signals of ability becomes less important as more information about productivity becomes available; this mirrors traditional insights about labor market productivity, in which college GPAs and degrees become less important as more information about a worker's productivity is released.

Our findings suggest that having the ability to signal productivity would not lead to appreciable differences in the amount of education that individuals choose to pursue. Within our sample, only a small fraction of ranked participants (less than seven percent) choose to enter the draft; furthermore, of those individuals, a vast majority are the most highly ranked individuals in our sample. Even among otherwise-exchangeable players, having an additional strong signal of productivity is an essential component of the entry decision, and players who do not possess the signal are very unlikely to enter the draft. In short, even among those deemed to be highly able, only those most able labor market entrants make the decision to forgo schooling to pursue a professional career immediately.

2 Background

2.1 Ability, Education, and Labor Market Outcomes

It is well known that individuals pursue education because of the pecuniary rewards associated with doing so. This pursuit may occur regardless of whether any additional human capital accumulation takes place, as described in the classic labor market signaling model formulated by Spence (1973). Despite the pecuniary incentives for additional years of schooling, individuals face uncertainty about the set of outcomes that could arise both from starting college and from continuing college in each subsequent year. Altonji (1993) [1] points out that there are large disparities between *ex ante* and *ex post* returns to starting college, and that these disparities vary appreciably by ability, as measured by SAT scores. Low-ability males are nearly twice as likely to drop out when compared to high-ability males and the *ex post* returns of males who attend fewer than two years of college are negative.

These findings indicate that labor market participants may be unaware of their own ability when entering college. As a consequence, college can serve as a tool for individuals to learn about their abilities and preferences from course grades, subject content, and information about future opportunities. Arcidiacano (2004)[2] explores this in the context of major choice and ability sorting and finds that a large fraction of sorting between majors occurs due to preferences for the subject matter. High grades serve as positive signals of ability, which facilitate staying in school and increase the likelihood of more math-intensive majors, such as those in natural science. Stange (2012)[27] adopts a similar approach to estimate the option value of continuing with one's college education. He finds that the option value to remain in school accounts for roughly 14 percent of the total value of enrollment, indicating that individuals benefit greatly from having the option to enroll or drop out. As in Arcidiacano (2004)[2], Stange (2012)[27] also finds that individuals use grades as barometers for whether to continue education. Trachter (2015)[28] extends this type of structural analysis to the decision of whether or not to pursue a four-year degree after a two-year degree. He finds that the ability to drop out of a two-year college explains 31 percent of the return to enrolling in a two-year college and that the ability to transfer to a four-year college explains the remaining 69 percent of returns; all of the returns are localized to learning about one's ability.

We examine the schooling decisions of individuals who possess external ratings of ability in a highly specialized labor market prior to pursuing college education. Employers (teams) are able to perfectly observe these ratings prior to deciding whether or not to hire (draft) employees and are willing to hire workers (players) that do not graduate from college based on their expected future productivity. Therefore, we observe how employers would hire individuals that do not complete college if they were able to observe credible signals of ability. More specifically, we examine the decisions that labor market participants make when their ability is already revealed to potential employers.

Most labor market participants cannot credibly reveal ability to their employers prior to college, and this problem at least partially drives their choices to pursue further education. Arcidiacono, Bayer and Hizmo (2010) [3] find that college graduates realize the returns to their own ability (as measured by AFQT scores) almost immediately upon entry into the labor force and that these returns do not change appreciably over the first 12 years of their careers. By contrast, however, they find that high school graduates who directly enter the labor force realize almost none of the returns to their own ability. Only after extensive labor market experience do high school graduates receive wages corresponding to their measures of ability. Fang (2006)[10] uses a structural model of education choice to decompose the college wage premium into productivity enhancement and ability signaling and finds that around one-third of the college wage premium could be attributed to ability signaling.

Our empirical setting varies from traditional education choices because high- and moderate-ability high school basketball players possess external ratings of own ability prior to entering college. These players may nonetheless extract additional information from their college performance about their abilities when deciding whether to leave college for the professional ranks. These additional signals are informative since professional teams only employ a small fraction of college basketball players, even if they are highly able, implying a considerable amount of uncertainty. Furthermore, because previous draft outcomes of individuals with similar abilities are observable, the marginal schooling decisions of these highly able labor market participants provide implicit information about these individuals' expected labor market outcomes associated with leaving or continuing in school. In other words, individuals in our setting are simultaneously learning about their own ability and the ability of other potential labor market entrants.

High school basketball players possess strong external signals of ability prior to entering college, but the decision of which college to attend is still important. Attending a college with better coaching and a history of strong basketball graduates can enhance a player's signal of ability to professional teams. As an example, a player who attends the University of Kentucky, Duke University, or the University of Kansas may be selected highly because the strong signals associated with being recruited by those universities convey ability to future employers. Analogously, researchers have recently examined which components of college most precisely contribute to earnings growth, specifically the extent to which where one attends college contributes to the college wage premium. Using longitudinal data from Colombia, MacLeod, Riehl, Saavedra and Urquiola (2015)[21] find that the reputation and identity of a college affect both initial earnings and the subsequent growth paths of a graduate, even after controlling for ability from admissions tests. Hoekstra (2009)[16] finds that males accepted to a state's flagship university received nearly 20 percent earnings when compared with similarly able males who were not admitted. Hastings, Neilson and Zimmerman (2013)[14] use Chilean longitudinal data to examine how the joint major/college selection affects lifetime earnings. Since students in Chile apply to a major and university simultaneously and are matched to a degree/university pair, the authors are able to utilize the variation around the match cutoffs to find large earnings effects for selective majors at the cutoff, averaging up to 9.1 percent of sample earnings.

2.2 Asymmetric Information, Ability Signals, and the High School Basketball

There is a large information asymmetry between the amount of information a high school basketball player may have (or believe that he has) about his own ability and the amount of information about his ability that he can credibly convey to a third party based on his in-game performance alone. Traditional measures of performance, such as the number of points that a player scores, are noisy signals of ability due to large differences in opponent quality across players making direct comparisons difficult. Basketball recruiting ratings agencies, such as Rivals or Scout, collect an extensive amount of information on player ability, including in-game performances and external assessments, in order to rate and rank high school basketball players. Of the 261 domestic basketball players from the 2002-2012 high school cohorts that have been drafted into the NBA in the first round between 2002 and 2015, only seven were not ranked by at least one of the recruiting agencies.

To overcome this paucity of reliable information, promising players take a number of measures to improve the quality of information that is available for recruiters and for rating agencies. First, players may attempt to enroll in private high schools where appreciable resources are spent on both fielding a high-quality basketball team and on seeking out similar high schools for games.[1] Players often participate in out-of-season camps to hone their skills and make themselves more noticeable to colleges and universities.[2] Evaluators from rating agencies can attend the camps and gather information about how highly-skilled high school players perform against each other, and these evaluators can combine information about the levels of competition within these camps with the information of how these players perform in high school to form the ratings that are disseminated to both subscribers and the public at large, which includes universities, NBA teams, and other interested parties.

While high school players can take steps to more clearly signal their ability to rating agencies, colleges and universities, and professional teams, they are unable to actually manipulate their ratings for a few reasons. One reason is that player performance within camps and games is observable. While direct comparisons between different levels of high school competition may be difficult, it is appreciably *less* difficult to compare players playing directly against each other within the same camp. Inaccurate representations of ability are not credible. Second, an offer of money from a player to a third party would be an NCAA violation, and would both compromise the player's ability to play basketball in college and serve as a negative signal to professional teams. Finally, ratings agencies profit directly from the quality of information that they provide to subscribers that expect rating agencies to offer impartial, unbiased assessments of player quality such that there is little incentive for an agency to compromise this notion for any particular player (Bricker and Hanson, 2013) [8].

Rating agencies use the information that they gain from high school games and from basketball camps to form two sets of ordinal rankings of all high school basketball players in the United States. The first set of ordinal rankings is at the individual player level. Because it is prohibitive to rank every single high school player in the country and because the differences in ability between players of moderate ability level may be small, agencies only rank construct ordinal rankings for

[1] These high schools, such as Oak Hill Academy or IMG Academy, often command high tuition and their games are routinely seen on cable television stations, such as ESPN, ESPN2, or ESPNU.

[2] Examples of these camps include the LeBron James Skills Academy, Five Star Basketball Camp, and Nike Sports Camps, among others.

the 100 (Scout) to 150 (Rivals) of high school players that they consider to be the highest ability levels.These ordinal rankings, summarized in Table 1 for high school graduating cohorts between 2002 and 2012, differ by rating agency and do not take into account factors such as the position that the player plays on the court.

The next set of ordinal rankings that rating agencies construct, also summarized by graduation cohort in Table 1, is less granular and includes a larger number of players each year. This "star rating" system classifies players according to the rating agencies' perceptions of player ability into different "star" categories. The most promising players are "five star" recruits and are considered to be markedly better than those players who are grouped into the "four star" category, who, in turn, are considered superior to "three star" or "two star" individuals. "Five star" recruits are informally thought of as the most likely to be professional prospects and usually have their choice of which university to attend. Table 1 illustrates that for both ratings agencies, this category typically consists of the top 25 players from the more disaggregated ordinal ranking, but the exact number of players considered to be "five star" varies from year to year.

"Four star" recruits are also considered to be very promising, and are thought to have the potential to be professional prospects, but are considered to be less able than "five star" recruits. "Four star" recruits often have their selection of many different universities to attend, albeit perhaps not every university. There is more variation across ratings agencies in the number of players that they consider to be "four stars" with Rivals averaging around 70 players each year and Scout averaging around 100 players each year with the designation. Finally, "three star" and "two star" players are considered to be lower ability relative to their "five star" and "four star" peers and both agencies classify a larger number of players each year into these categories increasing the noise in these signals. In summary, high school basketball players in our sample can receive up to four ordinal rankings, two that are more granular and two that are aggregated into a "good, better, best" type of ranking, depending upon whether they are ranked by one or both agencies. Our empirical analysis exploits discontinuities in these rankings as individuals that are similarly ranked at the granular level can be classified into different "star" categories at the threshold.

This asymmetric information problem faced by high school basketball players mirrors the problem that high school students and parents face when trying to signal ability to various colleges or universities. High school grade point averages may vary wildly between schools because of differing curricula, and even the number of activities in which a student participates may not be suitably comparable between schools. (For example, one could think that the requirements to be a participating member of a club in a less-competitive high school are potentially less stringent than those of a more-competitive high school or a high school with many more students). Because of this, special emphasis is placed on performing well on standardized tests such as the PSAT, ACT, or SAT, which are direct measures of performance that can be compared between students. As an example, performance on the PSAT can lead to being recognized as a National Merit Scholar, National Merit Finalist, or National Merit Semifinalist; these categories are analogous to star ratings within high school basketball. Students who perform poorly on these tests may be rated more poorly by admissions committees, regardless of their high school accomplishments.

2.3 The Professional Basketball Labor Market

Professional basketball is organized very similarly to other markets, such as medical residencies, in which new entrants into the labor market cannot directly choose which employer they will work for. A player that wants to play professional basketball in the NBA typically enters via the NBA Draft[3], which is a modified reverse-order entry draft in which the team that finished the previous season with the worst record is most likely to have the first choice of new entrants in the upcoming season.[4] The NBA Draft consists of two rounds of selections, and, absent any trades, each of the thirty teams is given one selection per round based on their record and the results of the weighted lottery.[5] Players that are interested in entering the NBA Draft can do so in one of two ways: either exhaust all of their collegiate eligibility[6] or forgo their remaining eligibility and formally declare

[3]Although the majority of NBA players (426 out of 572 non-foreign born players) enter the league via the NBA draft, it is possible for undrafted players to obtain contracts as free agents. Players that are undrafted play in fewer games, play fewer minutes per game, play fewer seasons, and earn less career earnings on average relative to their drafted counterparts. Additionally, being drafted is no guarantee of making an NBA team as 67 out of 493 non-foreign born draftees never play in an NBA game.

[4]To prevent teams from intentionally losing to obtain higher selections, the first fourteen slots are determined by a weighted lottery, in which teams with poorer records have a greater chance of winning.

[5]Draft pick spots, as well as the rights to sign a player, are frequently traded between teams.

[6]Collegiate basketball players are eligible to play in college for four seasons with exceptions commonly being granted to either develop additional skills, for medical reasons, or for transfers between schools.

their intentions to enter the NBA Draft following the end of the collegiate basketball season in April as "early entry" candidates.

Players that have exhausted their four years of eligibility are automatically eligible for the draft without any additional entry requirements. Players that declare early entry for the NBA Draft can withdraw their entry and retain their college eligibility if they do so by a well-defined deadline, typically set about one month following the declaration deadline. A player that declares early entry for the NBA Draft without withdrawing is not eligible to return to competitive collegiate basketball, but is able to re-enroll to finish his degree or to compete professionally in a different league. Early entry decisions, as well as draft outcomes, by graduation cohort and college class year are summarized in Table 2. Because of the institutional features of the NBA Draft and the NCAA policy which stipulates that a player who enters the draft may not continue to play in college, a player's choice to enter the draft is a specific decision to stop accumulating any more basketball-specific human capital by playing in college and to instead become a professional. Unlike other settings, a player cannot "go back" to college to pursue basketball-specific human capital once he chooses to enter the draft.

Starting with the 1999 Collective Bargaining Agreement (CBA) between players and the league, players that enter the NBA via the draft are only eligible to receive contracts that are guaranteed for multiple seasons *regardless of performance* if they are selected in the first round (NBPA, 1999)[23]. All other players are eligible to sign shorter term contracts, for less money, at the team's discretion. This feature of the draft adds additional uncertainty for any player who wishes to become a professional as there is no guarantee that a player who is not selected sufficiently high will employed by an NBA team in the upcoming season. Groothius, Hill, and Perri (2007) [13] identify that the 1999 CBA, which guaranteed that younger players would be kept on their lower salary, initial contract for longer lengths of time, incentivized franchises to draft relatively younger players, *ceteris parabis*, leading to a further unraveling of the matching market first identified in Li and Rosen (1998) [20].

Prior to the signing of the 2005 CBA, a high school basketball player could declare himself eligible for the NBA Draft without any additional entry restrictions.[7] Following the 2005 CBA

[7]Indeed, highly prominent players such as Kobe Bryant and LeBron James chose to bypass college attendance entirely in order to play professionally.

and first implemented for the 2006 NBA Draft, high school basketball players are required to be at least one year removed from high school and at least 19 years of age prior to being eligible for early entry (NBPA, 2005)[24]. Because of this change in policy, nearly every high school basketball player that was ranked in the graduating cohorts between 2006 and 2012 choose to attend a college or university for at least one year.

3 Data Description

The primary benefit of utilizing professional sports to examine labor markets is the amount and detail of available information regarding worker, firm, and co-worker characteristics, job performance, and labor market outcomes. We collect data from two separate ratings agencies–Rivals.com (Rivals) associated with Yahoo! Sports and Scout.com (Scout) associated with FoxSports.com– that rank high school basketball players based on perceived ability. In addition to the measures of ability, these agencies also include other observable characteristics such as player height and weight. Although we focus exclusively on the final posted rankings, most of the top high school basketball players are initially ranked as freshman or sophomores, with the ranking agencies updating their information and their scouting reports as these players progress through high school. As summarized in Table 1, we observe rankings data on the highest ability recruits for both agencies for all graduating classes between 2003 and 2012 with the exception of Scout in 2004. We observe rankings according to the star classification for both agencies for all graduating classes between 2002 and 2012.

The rankings data provide us with metrics of player ability prior to entering collegiate or professional basketball. We match these data with player performance statistics, at both the collegiate and professional basketball level, available from Sports-Reference.com for the 2002-03 through 2014-15 basketball seasons. Our sample of collegiate basketball players consists of all players who graduated high school between 2002 and 2012. The sample of professional basketball players consists of all players who graduated high school between 2002 and 2012, regardless of whether they attended college.[8] The inclusion of collegiate-level performance data allows us to determine whether the

[8]We exclude foreign-born professional players even if their date of birth would place them in the same high school

prospect rankings are sufficient signals of ability while controlling for observable, in-game measures of performance in determining a player's entry decision and draft position.

We consider traditional statistics, both total and per game, measuring player performance including points, rebounds, assists, steals, blocks, field goal percentage, three-point field goal percentage, free throw percentage, games played, minutes played, personal fouls, and turnovers. Furthermore, we also consider a handful of common advanced statistics designed to more accurately capture player contributions towards winning including PER, usage rate, effective field goal percentage, true shooting percentage, win shares, and win shares per 48 minutes of playing time. Table 3 presents the summary statistics for the college data. Berri, Brook, and Fenn (2011)[5] find that in addition to the performance metrics which impact draft position, height for position and team performance in the college basketball playoffs are also correlated with draft outcomes. We also control for other publicly available signals of player ability provided by third parties; namely, whether a player was selected as a McDonald's All-American in high school or as an AP All-American in college.

Our sample, summarized in Table 4, consists of 18,819 unique basketball players that were either ranked in high school, played collegiate or professional basketball, or entered the NBA Draft and were in high school graduating classes between 2002 and 2012. From this sample, 803 players (4.27%) declared for early entry into the NBA Draft and 652 players (3.46%) were drafted. In our empirical estimations, we drop 349 foreign-born players (1.85%) that declared for the NBA Draft without being ranked in high school or playing collegiate basketball. Of the remaining players in the sample, 7,028 players (38.05%) received a ranking in high school with the majority attending college for at least one season. Less than one percent (101 players) of our remaining sample did not attend college with 67 of those players being ranked in high school. All of the players that were drafted into the NBA without collegiate experience were ranked in high school. The likelihood of a collegiate basketball player being drafted into the NBA is small with only 2.55% (468 players) of collegiate basketball players in our sample being drafted between 2002 and 2015. However, the majority of collegiate players that entered the NBA Draft early (92.82%) or were drafted (94.23%)

cohorts.

were ranked in high school despite these players constituting only 37.90% of all college basketball players.

4 Empirical Justification

We begin by tracing out the choice faced by a player at time 0. The player does not know the exact probability of being drafted, but he forms belief $p(s)$ which corresponds to the probability of being selected in draft slot s. His beliefs are a function of his observable rankings θ, his perceived ability A, and his observable college productivity at time 0, X_0. Should he be selected in slot s, he will receive wage w_s, which is the collectively-bargained outcome associated with slot s. A player holds beliefs over all possible slots $s \in S$, such that $\sum_{s \in S} p(s) \leq 1$. With the remaining probabilistic belief the player will go undrafted and receive outside option w^*. Moreover, the player also receives the discounted sum of future earnings from basketball, which are a function of his draft slot, beliefs about ability, and signal of ability. To summarize, at time zero, a player believes that, if he enters the draft, he will earn:

$$\sum_{s \in S} p(s|\theta, A, X_0) \left[w_{s,0} + \sum_{t=1}^{T} \beta^t w_t(\theta, A, S) \right] + \left(1 - \sum_{s \in S} p(s|\theta, A, X_0) \right) \sum_{t=0}^{T} \beta^t w^* \tag{1}$$

If the player decides not to enter the draft, he will not be compensated at time 0; rather he will play another year in college and have unknown productivity X_1. The player does not know what his productivity will be, so his belief about future productivity is captured by $\mathbb{E}(X_1)$. The expected value of waiting to enter until the next year is:

$$\sum_{s \in S} p(s|\theta, A, \mathbb{E}(X_1), X_0) \left[w_{s,1} + \sum_{t=2}^{T} \beta^t w_t(\theta, A, S) \right] + \left(1 - \sum_{s \in S} p(s|\theta, A, \mathbb{E}(X_1), X_0) \right) \sum_{t=1}^{T} \beta^t w^* \tag{2}$$

In equilibrium, for a given set of beliefs $p()$, a player will enter the draft if (1) exceeds (2), will not enter the draft if (2) exceeds (1), and will be indifferent between entry if (1) is equal to (2). Furthermore, because ability A and future productivity X_1 are unobserved at time 0, a player must rely on θ, the external signal of ability, and X_0, the observable productivity based on college performance at time 0 to form the beliefs needed to make the decision about whether or not to enter the draft.

The key dynamics of the entry decision are as follows. First, since $w(s)$ is decreasing in s, i.e., a player's collectively-bargained wage goes up as their pick slot s goes down, (the player who is selected second earns a higher salary than the player who is selected with pick 60), beliefs which place high probability on low slots will induce entry into the draft. Next, if a player believes that his college productivity in the next season, $\mathbb{E}(X_1)$, will cause his expected slot to become worse, i.e., $p(s|\theta, A, \mathbb{E}(X_1), X_0)$ will increase for worse slots and will decrease for better slots, the player will enter the draft at time 0. Finally, and most importantly for our design, for a fixed ability A and productivity X_0, higher signals of ability θ will cause beliefs $p(s)$ to increase for better slots, i.e., as s falls, $\frac{\partial p(s)}{\partial \theta} > 0$.

We capitalize on the notion that better signals will increase players' beliefs that they will be drafted highly and will induce entry. We concentrate our empirical design on a region in the signal space where productivity is fixed, but signals of ability vary discontinuously. In particular, players are exchangeable based on their productivity (they have similar college statistics) and granular rankings of ability. However, some otherwise exchangeable players have appreciably better aggregated signals of ability, as they are categorized into different star levels. We are thus able to isolate the effect of an improved signal of ability on the decision to enter the labor market.

4.1 Descriptive Results

We examine the information content of the high school rankings via descriptive analysis of the draft and labor market outcomes for the cross-section of players that were either ranked in high school, played collegiate basketball, or both. Utilizing both the continuous relative ranking of ability as well as the discrete categorical rankings, we explore the information content of each ranking separately as well as the possibility of interactions between the different rankings. Following the player cross-section analysis, we descriptively estimate both the probability of being drafted using the player-college data for each collegiate class (i.e., Freshman, Sophomore, Junior, and Senior) separately before proceeding to estimating Cox proportional hazard models, with and without time-varying covariates, for early entry.

The descriptive results from the linear probability estimates on the player cross-section are summarized in Figure 1 for the probability that a player is drafted and in Figure 2 for the probability of playing in the NBA. For each set of rankings, we estimate:

$$Pr(outcome_i) = \beta_0 + \beta_1 rated_i + \beta_2 ranking_i + \mathbf{\Delta} star_i + \mathbf{\Gamma}(ranking_i * star_i) + \epsilon_i \qquad (3)$$

where $outcome_i$ is either being drafted or playing in the NBA, $rated_i$ is an indicator variable for whether the player was rated by the agency, $ranking_i$ corresponds to the numerical ranking of the player, and $star_i$ is a vector of indicator variables corresponding to the "star" categories. Figures 1 and 2 include both the fitted linear estimation as well as the raw average probabilities for players rated within a five-unit range (i.e., players ranked 1 to 5 are included in the first bin, players ranked 6 to 10 are included in the second bin, etc.). The fitted estimations assume that the top 25 rated players also receive five stars and the 26th to 100th rated players receive four stars. Regardless of the ranking agency or the outcome metric used, both figures suggest that the ranking and star ratings convey information regarding player ability with higher-ranked players and players with a higher star rating being more likely to both be drafted and to play in the NBA. Moreover, the statistically significant coefficient on the interaction between the ranking and the star rating imply that these different indicators may be either complementary or reinforcing.

We further explore the information content of high school rankings by examining the labor market entry choices of collegiate players as a function of prior ratings. For each class year and each ratings agency, we estimate the linear probability that a collegiate basketball player declares early entry into the NBA draft according to equation (3). The results of these descriptive estimations, reported in Table 5, reveal that players receiving a five- or four-star rating are significantly more likely to declare early entry into the NBA draft relative to their three-star, two-star, and unranked counterparts. The estimated magnitudes for five-star athletes remain relatively consistent across class years whereas juniors are more likely to declare for early entry if they receive a four- or three-star rating relative to sophomores who, in turn, are also more likely to declare relative to freshman receiving the same ratings. This is consistent with relative lower ability individuals delaying entry into the labor market in order to acquire additional years of human capital. Moreover, the interaction between the top ranking and a five-star rating is only significant for freshman players whereas the interaction between the top ranking and a four-star rating increases (in absolute magnitude) with each subsequent year of experience.

Our final descriptive empirical exercise is to further explore the potential dynamics underlying the entry decision by estimating Cox proportional hazard models with and without allowing the rankings variables to be time variant. The hazard rate $h(l)$ is the likelihood that a player enters the draft as a freshman, sophomore, or junior conditional on him not previously entering the draft. The estimated model using only the high school ratings and rankings is:

$$\begin{aligned} &h(l|rated_i, ranking_i, star_i, (ranking_i \times star_i)) = \\ h_0(l) \exp &(\beta_0 + \beta_1 rated_i + \beta_2 ranking_i + \delta star_i + \gamma(ranking_i \times star_i) + \epsilon_i) \end{aligned} \tag{4}$$

where the variables are as defined previously. The results from the proportional hazard models, reported in Table 6, confirm the significant and positive correlation between receiving a higher star rating and entering the NBA draft early. The time-varying estimates also confirm that the information contained in the five-star rating becomes less relevant for entry decisions across time. The results also support the complementarity between the receiving a five-star rating in addition to the top 150 or top 100 ranking. The interaction remains significant and negative, suggesting that individuals with more positive signals are more likely to enter early. However, the significant positive coefficient on the interaction term in the time-varying estimations implies that this relationship is strongest initially and weakening over time, a result consistent with the insignificant coefficient on the interaction terms for Sophomores and Juniors in the linear probability model for early entry.

5 Methodology: Regression Probability Jump and Kink (RPJK) Design

Our main analysis examines the change in the probability of entering the NBA draft around a threshold of players ranked above or below the 25th-best player in the Scout and Rivals rankings.

At this threshold, players have similar granular rankings, but differing star levels, as can be seen in Table 1. In the Scout rankings, 7 of 11 years have 25 or fewer five-star players, while in the Rivals rankings, 5 of 11 years have 25 or fewer five-star ratings. For this reason, we use 25 as the threshold for belief treatment in our data. Put simply, players ranked 25 or below are much more likely to be assigned five-star ratings, and players ranked higher than 25 are much more likely to be assigned four-star ratings.

5.1 Graphical Evidence for the RPJK Design

Because of the imprecise assignment rule, we are unable to implement a sharp regression-discontinuity design. In particular, for some years more than 25 players are ranked as "five-star," while in other years 25 or fewer players are ranked as "five-star." Furthermore, as can be seen in Figures 3 through 6, there is little evidence of a jump in the probability of entry at the threshold, indicating that we may not be able to achieve identification of a local average treatment effect solely from a fuzzy regression discontinuity design.[9] However, Figures 3 through 6 strongly suggest the presence of a kink in the probability of NBA Draft entry based on crossing the threshold. Figures 3 and 4 provide graphical evidence of the kink in the probability of entry based on going from four-stars to five stars for college freshmen. These players only have one year of collegiate productivity and, relative to sophomores or juniors, may need to rely more heavily on the quality of the signal in order to decide whether or not to enter the NBA draft. Figures 3 and 4 indicate that, prior to crossing the threshold, freshmen are systematically more likely to enter the draft. However, as they approach the threshold, the more granular signal becomes indicative of lower perceived ability, and as such the probability entering the draft falls. After crossing the threshold, the less granular signal becomes more likely to shift from five-star to four-star, and players after the threshold are unlikely to enter the draft at all.

Figures 5 and 6 showcase a similar kink. These figures include the full sample of freshmen, sophomores, and juniors. The slopes of entry probabilities below the threshold are less steep than the slopes in Figures 3 and 4, possibly because sophomores and juniors have less noisy estimates of

[9] In fact, Figures 3 and 5 would seem to indicate a jump in entry when going from five-star to four-star, which is inconsistent with the information content of such signals.

the expected benefit for playing in future seasons. The probability that players above the threshold enter the draft is larger than the probability for just freshmen, but is not appreciably higher than zero.

5.2 Estimation Strategy

Figures 3 through 6 lead us to employ a fuzzy regression probability jump and kink design, as described in Dong (2016) [9]. This is a flexible variant of the regression discontinuity design that allows for the possibility of both a jump and a kink, and is useful in instances when a kink in treatment is present, but the assignment rule is represented by a jump. We interpret our estimates of the kink as a local average treatment effect of the effect of being given a favorable aggregate ranking on the probability of entry. In particular, for a given player i with ranking x_i, we model the probability of obtaining a five-star rating 5_i^* as:

$$Pr(5_i^* = 1|x_i) = \begin{cases} g_1(x_i) & \text{if } x_i > 25 \\ g_0(x_i) & \text{if } x_i \leq 25 \end{cases} \tag{5}$$

where $g_1(x_i)$ and $g_0(x_i)$ are functions that guide the probability of a five-star ranking being assigned to a player at the cutoff. Next, define an indicator function T_i which takes the value 1 if $x_i \leq 25$ and 0 otherwise to identify whether or not a player is above the ranking threshold, or at or below the ranking threshold.

The running variable in our context is the player's ranking in either the Scout 100 or Rivals 150, x_i. We implement the fuzzy regression probability jump and kink design by using two-stage least squares; first, we instrument for the player's five-star status with the indicator function T_i by estimating:

$$\hat{5_i^*} = \delta_0 + \delta_1(x_i - 25) + \delta_2(x_i - 25) \times T_i + \rho T_i + \epsilon_{1,i} \tag{6}$$

after using the estimates from the first stage, we next estimate the probability of entry E for player i in the second stage:

$$E_i = \alpha+\beta_1(x_i-25)+\beta_2(x_i-25)\times T_i+\beta_3\left\{\delta_0+\delta_1(x_i-25)+\delta_2(x_i-25)\times T_i+\rho T_i+\epsilon_{1,i}\right\}+\epsilon_{2,i} \quad (7)$$

which is estimated via local linear regression. Following the recommendation of Gelman and Imbens (2014) [12], we avoid the use of higher-order polynomials in fitting our model, and as such report local linear results only. However, because players are unable to control their ranking and because each player is given an ordinal ranking (there is only one player per granular ranking per year), we are not concerned with manipulation of the running variable, as described in McCrary (2008) [22], or "heaping," in which a large number of players would be located immediately above the cutoff, as described in Barreca et al. (2016) [4].

The coefficients of interest in the model are β_2, which is the estimate of the kink in the slope of the probability of entry based on the interaction of the treatment (being ranked 25 or lower) with the distance to the threshold of 25, and β_3, which represents the estimate of the jump in the probability of entry when going from being a four-star to a five-star recruit. Players that are ranked more favorably have lower rankings, and as such as $(x_i - 25)$ is more negative. A negative slope on β_2 indicates that players who are ranked more highly around the cutoff are more likely to enter the draft, while a statistically significant coefficient on β_3 indicates that the probability of entry jumps when going from four-star to five-star. Based on the graphical evidence from Figures 3 through 6, we expect β_2 to be significant and negative, but do not expect to see a jump in the probability of entry.

6 Empirical Results

6.1 Establishing Exchangeability

Our estimates of the kink are compromised if other characteristics that would lead to entry also change at the discontinuity. Put simply, if players who are rated as being five-star have productivity characteristics which would cause them to enter the draft that four-star players do not have, then our estimates of the kink in productivity are biased upward and we do not achieve identification.

To investigate this possibility, we examine the difference in means for eight key productivity characteristics in our data: points per game, rebounds per game, steals per game, blocks per game, assists per game, field goals per game, field goal percentage, and minutes played. Each of these characteristics positively contribute to the perceived productivity of a player and would be considered part of X_0, the observable productivity of the player at the time the entry decision is made. If players who receive the five-star treatment have higher values of these productivity characteristics, they may be more inclined to enter the draft.

Tables 7 produces a balance table for Rivals rankings using our entire sample for two separate bandwidths: players ranked between 11-40, and players ranked between 16-35. The first panel of Table 7 shows the results for the wider bandwidth; the first column produces means and standard deviations of the productivity characteristics for players ranked greater than 25, while the second column produces means and standard deviations of those characteristics for those ranked 25 or below. The third column contains results of a t-test of unequal variances regarding the difference in means between the first and second columns. We find that there are large statistical differences in many of the productivity characteristics, including points, rebounds, field goals, and minutes per game, which as a whole indicates that these players should not be thought to be exchangeable. Estimates of a kink within this bandwidth are likely to be biased. This finding is unsurprising since the bandwidth is very large and players ranked very close to the top 10 are plausibly different from those closer to 41-50.

The second panel of Table 7 shows the results for the smaller bandwidth, where players are ranked between 16-35. We find much stronger evidence for balance within this particular bandwidth. In particular, we only find statistical differences for field goal percentage ($p < .05$), and rebounds per game at the ten-percent level ($p < .1$). Because field goal percentages may be based on vastly different numbers of shots, we do not consider the difference in field goal percentage to be especially troublesome, and we interpret the second panel as strong evidence of balance at the discontinuity. We also consider the second panel to be our preferred set of estimates due to the fact the players in this bandwidth should be considered exchangeable, while players in the first panel

Table 8 reproduces the balance table for the subsample of freshmen in our data set. Again, the first panel considers the bandwidth from 11-40, while the second panel considers the bandwidth from 16-35. The first panel demonstrates that the covariates do not achieve balance for the 11-40 bandwidth, as there are statistical differences for nearly all of the productivity characteristics. However, the second panel shows that, within the 16-35 bandwidth, productivity characteristics are balanced.

Tables 9 and 10 produce similar balance tables for the Scout rankings. Within the Scout rankings, we again see pronounced statistical differences in the 11-40 bandwidths for both the full sample and the subsample of freshmen, indicating a lack of balance of covariates. Furthermore, we see limited evidence for balance within the 16-35 bandwidth for the full sample, as there are pronounced statistical differences in blocks and rebounds per game. However, we do have evidence for balance within the subsample for freshmen.

As a consequence of our investigations of balance at the threshold, we note that our estimates of the kink for the bandwidth of 16-35 are most likely to satisfy the assumptions required for the RPJK design, and are most likely to do so for the Rivals subsample. While we report the estimates for alternative bandwidths and for the Scout data, we recognize that the assumptions need for the RPJK design may not hold. As such, our preferred estimates are for the bandwith of 16-35 and for the Rivals rankings.

6.2 Results

Tables 11 and 12 produce the main results. Table 11 produces estimates of both the jump in the probability of entry based on the movement from four-star to five-star, and the kink in the slope of the entry decision for players marginally above the threshold. All models are estimated using local linear specifications and are produced for two different samples (freshmen and the full sample) and two different bandwidths (11-40 and 16-35).

Table 11 produces the results for rankings by Rivals. Column 1 presents local linear estimates of the second stage regression for freshmen ranked between 11-40. We do not find evidence of a jump, as the coefficient on five star is insignificant. However, we do find evidence of a kink in the entry function. In particular, freshmen ranked between 11-25 are 2.8 percent more likely to enter the draft as their ranking approaches 1. We note that, within this sample, we did not find balance of the covariates, so our estimate contains omitted variable bias, as described in Lee and Lemieux (2010)[19]. Nonetheless, this is suggestive evidence of a kink in the entry function.

Column 2 presents estimates for the sample of freshmen ranked between 16-35. Recall that, within this sample, we found balance of the covariates. As such, omitting the covariates from our estimates does not lead to any omitted variable bias. Within this sample, we again do not find evidence of a jump in the probability of entry if a player is ranked with five stars. However, we find that a decrease in ranking by one slot for those treated as five stars increases the probability of entry by 3.8 percent relative to players receiving only four star ratings.

Columns 3 and 4 reproduce the analysis for the full sample. For these players, the impact of the signal on the decision to enter the draft may weaken as they have gained more information about their productivity from their college performance. Indeed, in both Column 3 (players ranked 11-40) and Column 4 (players ranked 16-35) we do not find any evidence of a jump in the probability of entry for being ranked five stars. However, we continue to find evidence of a kink; in particular, players ranked between 16-35 in the full sample exhibit a 2.7 percent higher chance of entering the draft for a one-slot fall in ranking if they are ranked as a five-star recruit. [10]

Our preferred estimates are in Columns 2 and 4, where we achieve balance around the threshold. In these estimates, we find that players categorized as being of higher ability are much more responsive to a marginal increase in their more granular signal of ability when compared to players categorized as being slightly less able, despite being otherwise exchangeable. These results are strongest for freshmen, who have little information on their own productivity from college performance, but are still pronounced for upperclassmen, who have more information on own productivity from playing additional seasons. More generally, our results suggest that only individuals with multiple strong external signals of ability make the decision to forgo additional schooling, even if the other information that these individuals possess regarding their productivity is the same.

As illustrated in Tables 9 and 10, we do not achieve balance at the discontinuity for the Scout

[10] We point out that the constant is significant and positive for players in our full sample ranked between 11-40 (Column 3). We attribute this to the fact that our sample does not exhibit balance within this threshold, and as such the constant includes the variation in entry for players that score more frequently, collect more rebounds, blocks, and steals, and shoot a higher percentage. All of these attributes increase the probability of entry.

data, and as such our estimates for the RPJK design are likely to lack identification. This can be seen in Table 12, which again estimates the jump in the probability of entry based on the movement from four-star to five-star and the kink in the slope of the entry decision for players marginally above the threshold. All models are again estimated using local linear specifications and are produced for two different samples (freshmen and the full sample) and two different bandwidths (11-40 and 16-35). We find evidence of a jump from five-star to four-star (which can be seen in Figures 3 and 5) and is inconsistent with the notion that players with stronger signals should enter more often. However, the magnitude of the jump is mitigated in all specifications by the positive and statistically significant constant, which captures the variation associated with the missing covariates that are unbalanced in the threshold. As a consequence of this lack of balance, estimates of the kink in entry are at best imprecise. We find limited evidence of a kink in Column 1 (Freshmen, ranked 11-40), but have imprecise estimates of the kink in Columns 2-4.

7 Conclusion

The choice to pursue additional years of schooling is routinely framed in the context of making a choice between signaling productivity and human capital accumulation. Because most prospective labor market entrants cannot credibly signal ability before entering college, it is not possible to consider what education choices they would have made had they possessed signals of ability. It is thus unclear what education choices people would make if they could credibly communicate their ability, nor is it clear how much ability they would have to convey in order to forgo additional schooling.

In this paper we examine a specific labor market where individuals credibly possess external signals of ability prior to making a schooling choice. In professional basketball, promising labor market entrants are often rated using two ordinal scales–one that is more granular, and one in the "good, better, best" style. We find that, conditional on observable productivity, possessing a granular signal of ability does not induce most college players to forgo additional schooling in order to enter the NBA Draft. However, players rated as having the best ability use the additional information conveyed in the granular ranking to inform their decision about whether or not to enter the draft. In particular, we estimate that a one-slot reduction in ranking (going from the 25th- to the 24th-best player, for example) increases the likelihood of entering the draft by anywhere from 1.6 to 3.8 percent per slot. These results are robust to both inexperienced and experienced college players.

Our evidence suggests that possessing signals of ability may not impact the level of schooling for all but the most able labor market participants. Indeed, within our sample, only 453 of the 7,000+ players enter the draft. However, among those players, players rated to be of the highest ability use their relative status among other high-ability players to assist in making the determination about whether or not to enter the draft *even after knowing their productivity.* Among the most able labor market participants, small differences in ability assessment may lead to appreciable and significant decreases in the amount of schooling pursued.

References

[1] Joseph Altonji. The Demand for and Return to Education When Education Outcomes Are Uncertain. *Journal of Labor Economics*, 11(1):48–83, 1993.

[2] Peter Arcidiacano. Ability Sorting and the Returns to College Major. *Journal of Econometrics*, 121:343–375, 2004.

[3] Peter Arcidiacano, Patrick Bayer, and Aurel Hizmo. Beyond Signaling and Human Capital: Education and the Revelation of Ability. *American Economic Journal: Applied Economics*, 2(1):76–104, 2010.

[4] Alan Barreca, Jason Lindo, and Glen Waddell. Heaping-Induced Bias in Regression-Discontinuity Designs. *Economic Inquiry*, 54(1):268–293.

[5] David Berri, Stacey Brook, and Aju Fenn. From College to the Pros: Predicting the NBA Amateur Player Draft. *Journal of Productivity Analysis*, 35(1):25–35, 2011.

[6] Lex Borghans, Bas ter Weel, and Bruce Weinberg. Interpersonal Styles and Labor Market Outcomes. *Journal of Human Resources*, 43(4):815–858, 2008.

[7] Samuel Bowles, Herbert Gintis, and Melissa Osborne. The Determinants of Earnings: A Behavioral Approach. *Journal of Economic Literature*, 39(4):1137–1176, 2001.

[8] Jesse Bricker and Andrew Hanson. The Impact of Early Commitment on Games Played: Evidence from College Football Recruiting. *Southern Economic Journal*, 79(4):971–983, 2013.

[9] Yingying Dong. Jump or Kink? Regression Probability Jump and Kink Design for Treatment Effect Evaluation. unpublished manuscript, UC-Irvine.

[10] Hanming Fang. Disentangling the College Wage Premium: Estimating a Model with Endogenous Education Choices. *International Economic Review*, 47(4):1151–1185, 2006.

[11] Alfonso Flores-Lagunes and Audrey Light. Interpreting Degree Effects in the Returns to Education. *Journal of Human Resources*, 45(2):439–467, 2010.

[12] Andrew Gelman and Guido Imbens. Why High-Order Polynomials Should Not be Used in Regression Discontinuity Designs. NBER WP 20405.

[13] Peter Groothuis, James Hill, and Timothy Perri. Early Entry in the NBA Draft: The Influence of Unraveling, Human Capital, and Option Value. *Journal of Sports Economics*, 8(3):223–243, 2007.

[14] Justine Hastings, Christopher Neilson, and Seth Zimmerman. Are Some Degrees Worth More Than Others? Evidence from College Admission Cutoffs in Chile. 2013. NBER WP 19241.

[15] Jerry Hausman and Gregory Leonard. Superstars in the National Basketball Association: Economic Value and Policy. *Journal of Labor Economics*, 15(4):586–624, 1997.

[16] Mark Hoekstra. The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach. *The Review of Economics and Statistics*, 91(4):717–724, 2009.

[17] Thomas Hungerford and Gary Solon. Sheepskin Effects in the Returns to Education. *The Review of Economics and Statistics*, 69(1):175–177, 1987.

[18] David Jaeger and Marianne Page. Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education. *The Review of Economics and Statistics*, 78(4):733–740, 1996.

[19] David Lee and Thomas Lemieux. Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48:281–355, 2010.

[20] Hao Li and Sherwin Rosen. Unraveling in Matching Markets. *American Economic Review*, 88(3):371–387, 1998.

[21] W. Bentley MacLeod, Evan Riehl, Juan Saavedra, and Miguel Urquiola. The Big Sort: College Reputation and Labor Market Outcomes. 2015. NBER WP 21230.

[22] Justin McCrary. Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics*, 142(2):698–714, 2008.

[23] National Basketball Players Association (NBPA). *1999 Collective Bargaining Agreement*, 1999.

[24] National Basketball Players Association (NBPA). *2005 Collective Bargaining Agreement*, 2005.

[25] Ellen Nyhus and Empar Pons. The Effects of Personality on Earnings. *Journal of Economic Psychology*, 26(3):363–384, 2005.

[26] Michael Spence. Job Market Signaling. *Quarterly Journal of Economics*, 87(3):355–374, 1973.

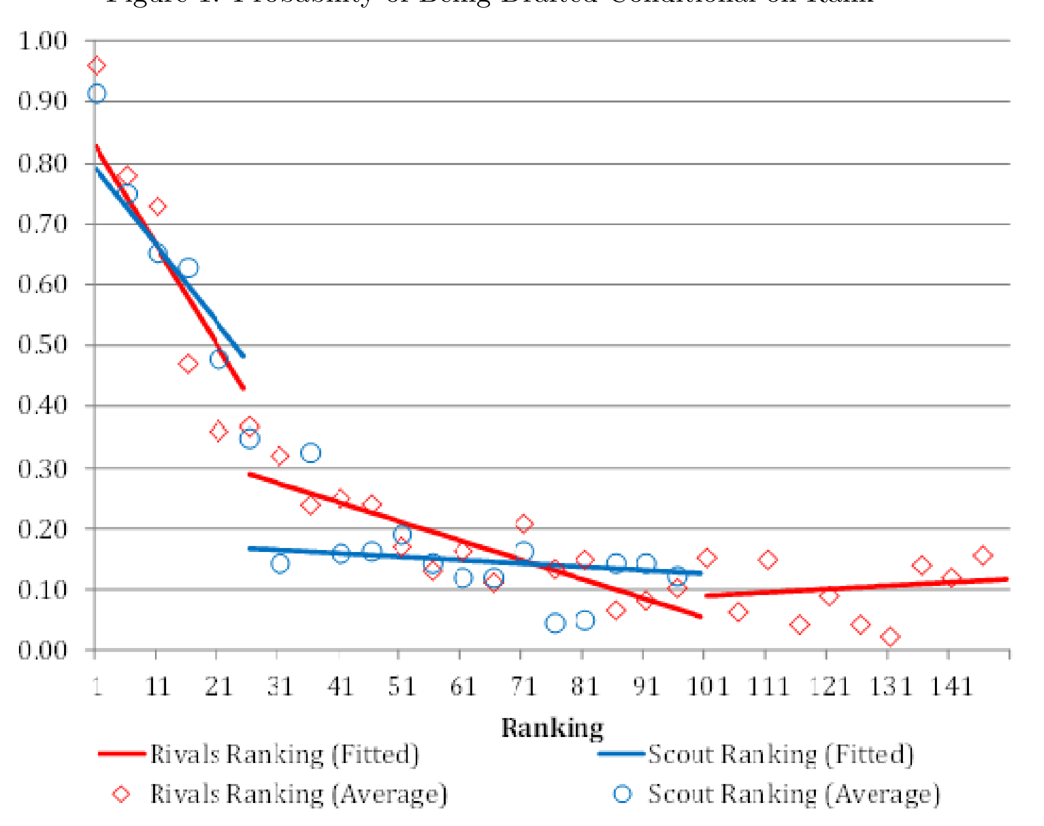
[27] Kevin Stange. An Empirical Investigation of the Option Value of College Enrollment. *American Economic Journal: Applied Economics*, 4(1):49–84, 2012.

[28] Nicholas Trachter. Stepping Stone and Option Value in a Model of Postsecondary Education. *Quantitative Economics*, 6:223–256, 2015.

[29] John Tyler, Richard Murname, and John Willett. Estimating the Labor Market Signaling Value of the GED. *Quarterly Journal of Economics*, 115(2):431–468, 2000.

[30] Andrew Weiss. Human Capital vs. Signaling Explanations of Wages. *Journal of Economic Perspectives*, 9(4):133–154, 1995.

Figure 1: Probability of Being Drafted Conditional on Rank



?//

Figure 2: Probability of Playing in the NBA Conditional on Rank

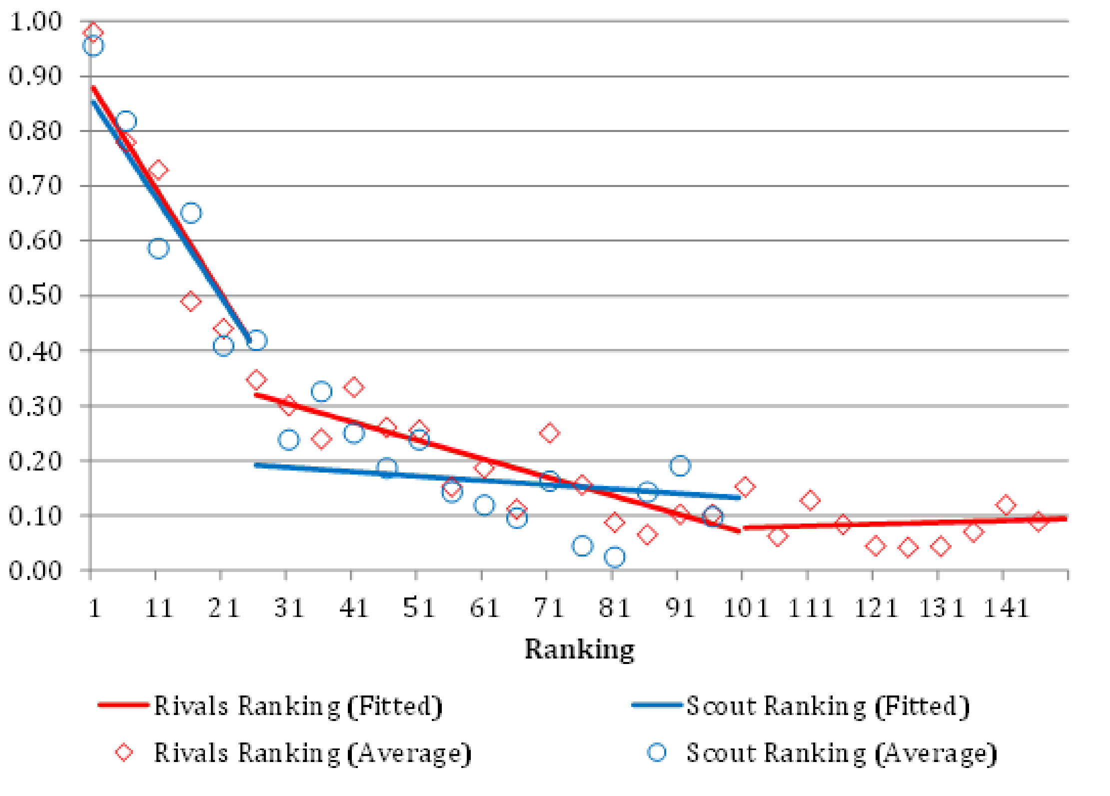


Figure 3: Entry Decisions of College Basketball Players, Freshmen Only, Scout Rankings

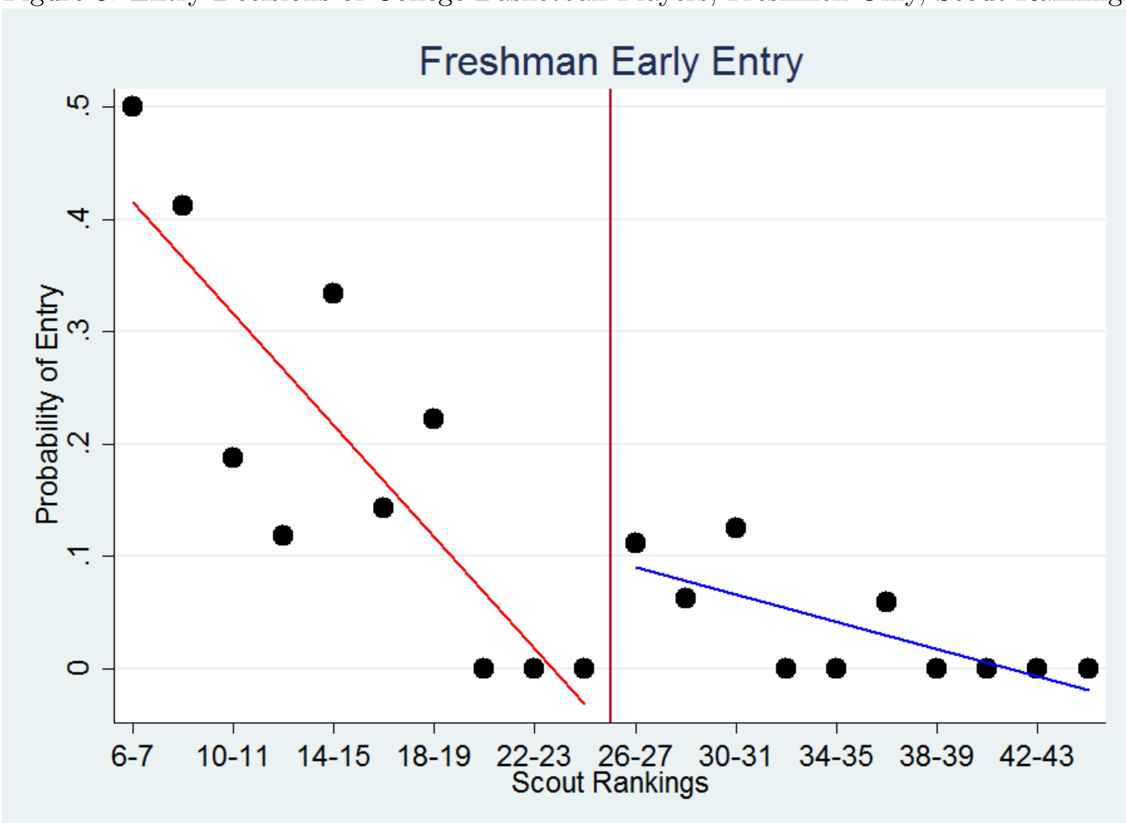


Figure 4: Entry Decisions of College Basketball Players, Freshmen Only, Rivals Rankings

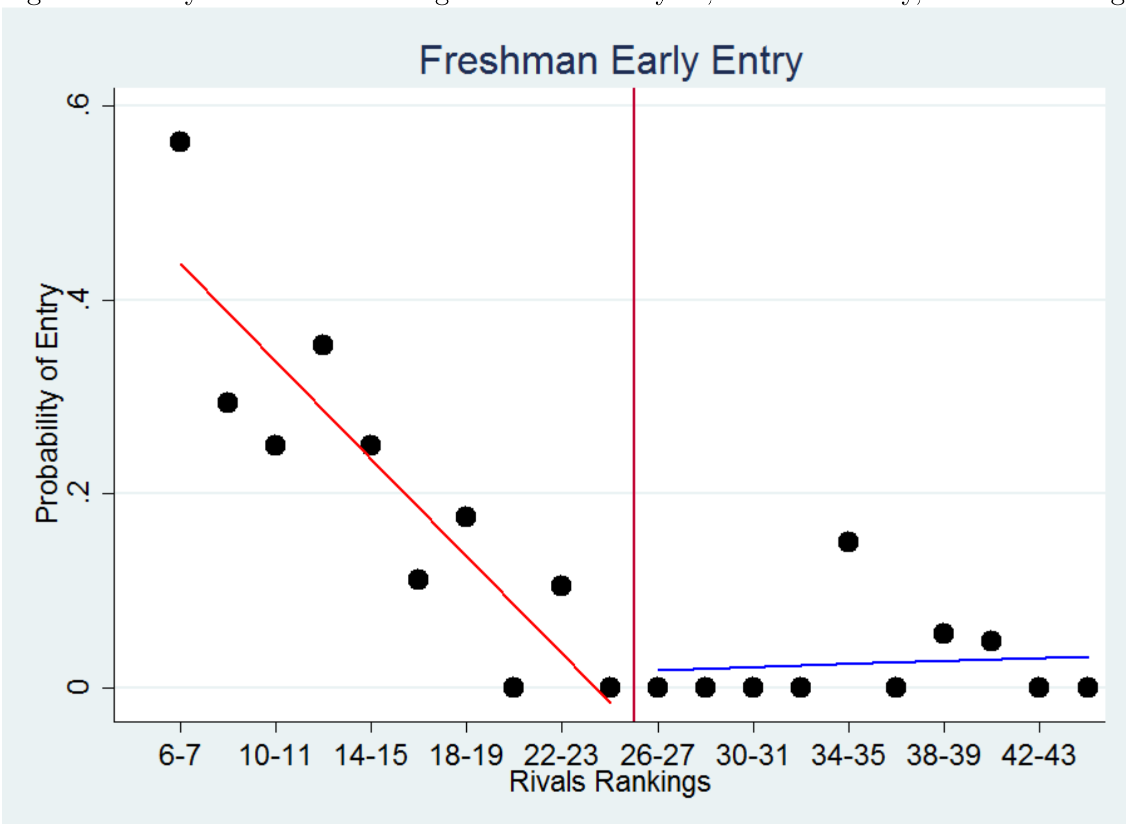


Figure 5: Entry Decisions of College Basketball Players, All Underclassmen, Scout Rankings

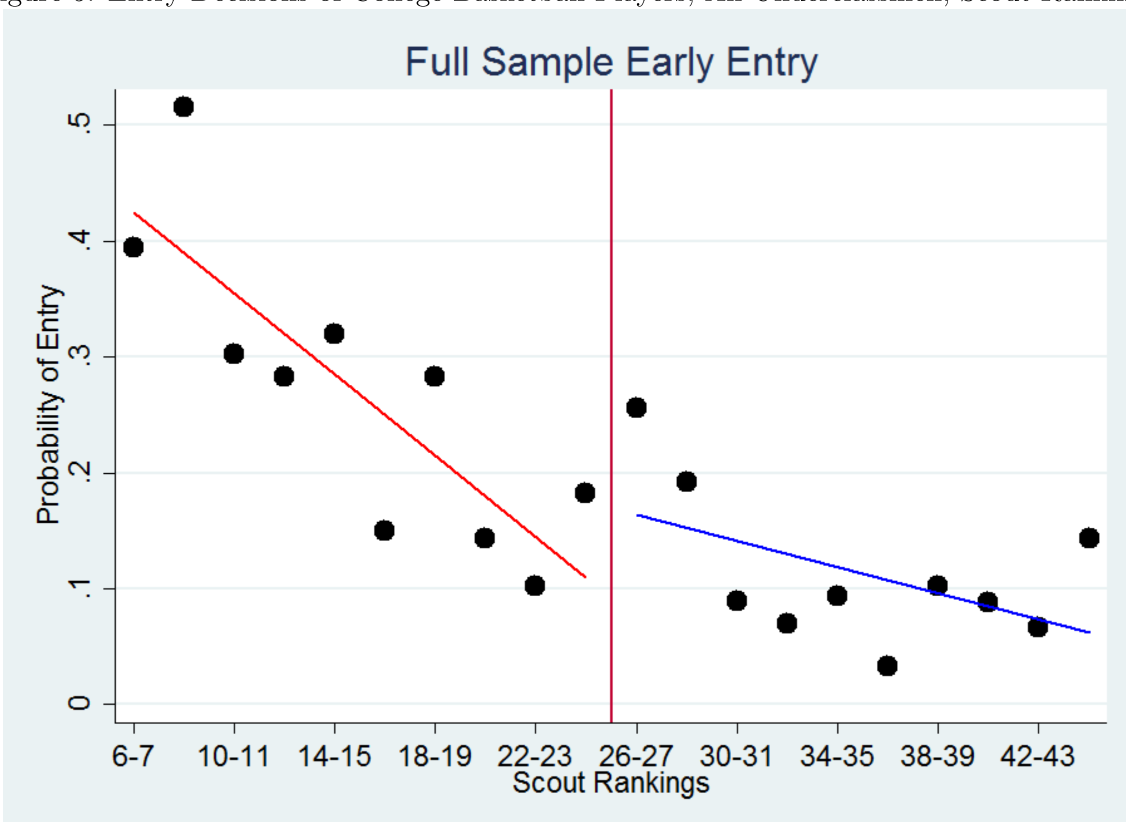


Figure 6: Entry Decisions of College Basketball Players, All Underclassmen, Rivals Rankings

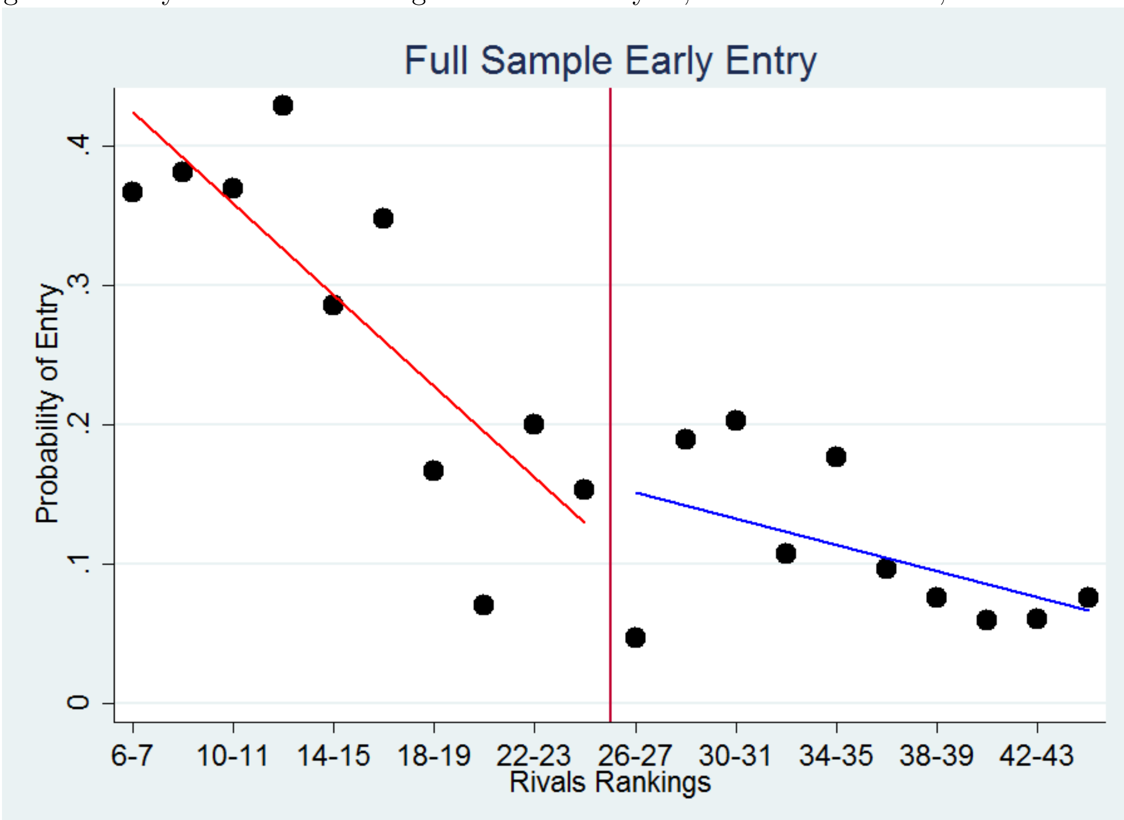


Table 1: Summary of High School Rankings

		Rivals Rankings					Scout Rankings				
Year	Observations	Top 150	5*	4*	3*	2* or 1*	Top 100	5*	4*	3*	2*
2002	502	1	18	33	133	45	1	25	116	116	136
2003	509	138	25	46	214	38	91	17	90	90	201
2004	800	147	32	54	161	30	6	26	90	90	519
2005	775	135	24	69	170	79	92	24	87	87	419
2006	686	139	29	69	203	54	94	27	106	106	297
2007	817	144	28	77	255	241	97	26	88	173	413
2008	768	139	26	66	239	271	91	24	83	140	393
2009	672	144	25	77	254	119	98	25	82	128	353
2010	548	146	28	71	217	53	94	25	80	142	249
2011	497	140	28	78	234	38	98	26	80	147	177
2012	456	141	25	79	222	44	100	26	82	152	125
All Years	7042	1415	288	719	2302	1012	862	271	984	1172	3282

Table 2: Summary of College Productivity Data

	Full Sample					Players Ranked in High School				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
College Productivity Variables										
Seasons Played	47,172	3.128	1.094	0	6	21,327	3.466	0.939	0	6
Games Played	47,070	24.391	9.692	1	41	21,251	27.275	8.145	1	41
Minutes Played	24,117	522.733	368.508	0	1544	10,784	635.274	351.080	0	1543
Points	47,070	166.639	157.739	0	1068	21,251	208.588	164.998	0	1068
Rebounds	47,070	76.086	67.890	0	508	21,251	93.712	70.200	0	504
Assists	47,070	31.850	37.240	0	351	21,251	39.432	40.371	0	351
Steals	47,070	16.229	15.677	0	125	21,251	19.614	16.107	0	121
Blocks	47,070	8.212	13.406	0	196	21,251	10.700	15.249	0	196
Turnovers	32,189	33.313	27.188	0	160	14,753	39.740	27.043	0	160
Personal Fouls	23,603	46.510	30.808	0	139	10,308	54.893	28.198	0	139
Field Goal Percentage	46,941	0.412	0.146	0	1	21,232	0.431	0.107	0	1
Three-point Field Goal Percentage	37,503	0.287	0.174	0	1	17,237	0.300	0.153	0	1
Free Throw Percentage	43,371	0.645	0.183	0	1	20,751	0.658	0.161	0	1
Advanced College Productivity Variables										
PER	23,568	12.249	9.672	-155.4	407.0	10,295	14.176	7.939	-53.5	407
Usage Rate	23,571	19.078	6.650	0	100	10,296	19.531	5.535	0	100
Effective Field Goal Percentage	46,940	0.461	0.159	0	1.5	21,232	0.479	0.107	0	1.5
True Shooting Percentage	46,959	0.490	0.147	0	1.5	21,236	0.510	0.099	0	1.5
Win Share	47,066	1.269	1.466	-2.3	11.3	21,249	1.692	1.599	-1.4	11.3
Player Characteristics										
Height in High School	20,974	77.014	3.458	64	88	20,925	77.017	3.458	64	88
Weight in High School	19,491	196.503	25.216	112	345	19,448	196.519	25.218	112	345
McDonald's All-American	47,131	0.145	0.119	0	1	21,319	0.032	0.175	0	1
AP All-American	47,064	0.007	0.127	0	3	21,251	0.014	0.183	0	3

Table 3: Summary of Early Entry Decisions

Year	Declare	Enter	Drafted	Declare	Enter	Drafted	Declare	Enter	Drafted	Declare	Enter	Drafted	Declare	Enter	Drafted
2002	13	7	5	3	3	1	-	-	-	-	-	-	-	-	-
2003	45	26	23	6	5	5	7	4	2	-	-	-	-	-	-
2004	63	27	20	13	9	8	5	3	3	3	1	1	-	-	-
2005	109	57	35	11	9	7	4	2	1	13	9	7	41	21	10
2006	93	46	32	1	-	-	5	3	2	13	10	9	37	19	13
2007	85	37	32	-	-	-	9	8	8	8	5	5	37	16	14
2008	90	44	34	-	-	-	13	13	12	15	10	9	32	14	8
2009	106	51	32	-	-	-	8	8	4	15	11	9	47	20	13
2010	113	72	33	-	-	-	15	10	9	20	15	10	48	31	11
2011	88	47	34	-	-	-	8	7	6	11	9	7	41	20	14
2012	68	57	35	-	-	-	10	10	10	16	16	13	21	20	8
2013	80	62	38	-	-	-	7	7	7	13	13	9	22	22	11
2014	58	42	24	-	-	-	-	-	-	15	15	11	18	17	8
2015	41	24	12	-	-	-	-	-	-	-	-	-	19	17	8
All Years	1,053	453	390	35	27	22	91	75	64	142	114	90	363	217	118

Table 4: Summary of Sample

Unique Players	Full Sample	Early NBA Entrants	Drafted	Played In NBA
Total	18,819	803	652	679
Foreign-born*	349	306	159	107
No College	101	79	25	28
Ranked in High School	67	48	25	25
Ranked by Rivals	52	41	24	24
Ranked by Scout	66	47	25	25
Played in College**	18,369	418	468	544
Ranked in High School	6,961	388	441	498
Ranked by Rivals	4,273	363	406	448
Ranked by Scout	6,252	378	437	486
1 Year Experience	4,775	89	67	67
2 Years Experience	4,674	121	100	115
3 Years Experience	2,997	195	125	130
4 Years Experience	5,670	13	171	228
5-6 Years Experience	253	0	5	4

*: classified as foreign in early entry or in NBA Draft data
**:

Table 5: Linear Probability Model (Early Entry)

	Freshman		Sophomore		Junior	
	Rivals	Scout	Rivals	Scout	Rivals	Scout
Rated by Agency?	0.002	0.001	0.001	0.004***	0.008	0.010***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.005)	(0.003)
Ranking	0.000***	0.000	0.000	0.000	0.000***	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
5 Star Rating Indicator	0.499***	0.437***	0.370***	0.371***	0.453***	0.503***
	(0.063)	(0.051)	(0.061)	(0.059)	(0.097)	(0.072)
4 Star Rating Indicator	0.040**	0.024***	0.125***	0.052***	0.352***	0.172***
	(0.016)	(0.007)	(0.026)	(0.012)	(0.043)	(0.020)
3 Star Rating Indicator	0.002	0.001	0.007**	0.003	0.027***	0.023***
	(0.002)	(0.001)	(0.003)	(0.003)	(0.007)	(0.006)
Ranking × 5 Star	-0.018***	-0.015***	-0.003	-0.003	-0.006	-0.008
	(0.004)	(0.003)	(0.003)	(0.003)	(0.006)	(0.004)
Ranking × 4 Star	0.000 **		-0.001***		-0.004***	
	(0.000)		(0.000)		(0.001)	
Constant	0.000**	0.001**	0.001***	0.001**	0.009***	0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
R-squared	0.267	0.234	0.164	0.156	0.117	0.108
N	12,777	12,777	11,184	11,184	13,099	13,099

Table 5 reports the linear probabilities of entry contingent on player ranking.
Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Cox Proportional Hazards Model (Early Entry)

	Time Invariant Estimates		Time-Varying Estimates			
			Rivals		Scout	
	Rivals	Scout	Main	Time-Varying	Main	Time-Varying
Rated by Agency?	0.214	1.069***	-0.945	0.430	0.492	0.209
	(0.399)	(0.218)	(2.039)	(0.743)	(0.829)	(0.324)
Ranking	0.003	-0.005**	-0.004	0.003	-0.006	0.001
	(0.002)	(0.002)	(0.007)	(0.003)	(0.007)	(0.003)
5 Star Rating Indicator	5.306***	4.399***	8.669***	-1.527**	7.123***	-1.251***
	(0.403)	(0.193)	(2.014)	(0.734)	(0.655)	(0.256)
4 Star Rating Indicator	4.022***	2.350***	4.191**	-0.067	3.173***	-0.333
	(0.405)	(0.193)	(2.087)	(0.537)	(0.716)	(0.268)
3 Star Rating Indicator	1.502***	0.802***	2.099	-0.228	0.936	-0.060
	(0.398)	(0.200)	(2.049)	(0.741)	(0.748)	(0.276)
Ranking $\times$ 5 Star	-0.051***	-0.030***	-0.101***	0.028***	-0.087***	0.029***
	(0.010)	(0.009)	(0.019)	(0.008)	(0.020)	(0.009)
Ranking $\times$ 4 Star	-0.021***	-	0.004	-0.009**		
	(0.003)		(0.013)	(0.005)		
Wald ($\chi^2(7)$)	1390.22	1358.63	1509.91		1436.12	
N	47,063	47,063	47,063		47,063	

Table 6 reports the hazard rates associated with estimates of entry; ”Failure” as defined in our model is when a player leaves college to enter the NBA draft.
Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 7: Balance of Covariates: Rivals Rankings

	Bandwidth: Rank 11-40			Bandwidth: Rank 16-35		
Variable	Rank > 25	Rank ≤ 25	Difference	Rank > 25	Rank ≤ 25	Difference
Points (per game)	9.359	10.565	-1.206***	9.893	9.945	-.052
	(4.994)	(4.636)	[-3.71]	(4.858)	(4.580)	[-0.13]
Rebounds (per game)	4.083	4.745	-.662***	4.286	4.615	-.329*
	(2.214)	(2.246)	[-4.38]	(2.237)	(2.208)	[-1.83]
Assists (per game)	1.746	1.858	-.113	1.778	1.708	.070
	(1.468)	(1.431)	[-1.15]	(1.452)	(1.345)	[0.61]
Blocks (per game)	.581	.649	-.068	.607	.630	-.023
	(.665)	(.614)	[-1.57]	(.649)	(.585)	[-0.47]
Steals (per game)	.868	.911	-.043	.860	.856	.004
	(.519)	(.501)	[-1.25]	(.485)	(.462)	[0.10]
Field Goals (per game)	3.303	3.772	-.469***	3.492	3.592	-.100
	(1.721)	(1.617)	[-4.15]	(1.686)	(1.639)	[-0.74]
Field Goal Percentage	.451	.469	-.018***	.456	.470	-.014**
	(.088)	(.075)	[-3.20]	(.077)	(.080)	[-2.18]
Minutes (per game)	24.841	26.301	-1.459**	25.320	25.259	.079
	(8.612)	(7.041)	[-2.19]	(8.341)	(7.154)	[0.03]
N	484	396	880	319	291	610

The first and second columns for each subsample contain means with standard deviations below. The third column contains t-statistics from a t-test of unequal variances on a difference in means in the covariates. Standard errors are in parentheses, while t-statistics are in brackets.
*** : $p < .01$, ** : $p < .05$, * : $p < .1$

Table 8: Balance of Covariates: Rivals Rankings, Freshmen Only

	Bandwidth: Rank 11-40			Bandwidth: Rank 16-35		
Variable	Rank > 25	Rank ≤ 25	Difference	Rank > 25	Rank ≤ 25	Difference
Points (per game)	6.751	8.899	-2.148***	7.330	7.957	-.627
	(4.174)	(4.316)	[-4.24]	(4.313)	(3.938)	[-1.05]
Rebounds (per game)	3.205	4.172	-.967***	3.429	3.840	-.411
	(1.849)	(2.074)	[-4.12]	(1.860)	(1.51)	[-1.83]
Assists (per game)	1.370	1.612	-.242	1.436	1.437	.001
	(1.260)	(1.394)	[-1.52]	(1.309)	(1.321)	[-0.01]
Blocks (per game)	.466	.591	-.125*	.491	.536	-.045
	(.521)	(.585)	[-1.89]	(.510)	(.494)	[-0.61]
Steals (per game)	.690	.845	-.155***	.714	.761	-.047
	(.457)	(.531)	[-2.62]	(.455)	(.487)	[-0.69]
Field Goals (per game)	2.411	3.171	-.760***	2.632	2.879	-.247
	(1.443)	(1.522)	[-4.30]	(1.499)	(1.446)	[-1.15]
Field Goal Percentage	.435	.458	-.023**	.439	.458	-.019
	(.098)	(.072)	[-2.33]	(.090)	(.076)	[-1.62]
Minutes (per game)	20.739	23.877	-3.138**	21.894	22.235	-.341
	(9.222)	(7.542)	[-2.30]	(8.571)	(7.574)	[-0.21]
N	146	136	282	98	92	190

The first and second columns for each subsample contain means with standard deviations below. The third column contains t-statistics from a t-test of unequal variances on a difference in means in the covariates. Standard errors are in parentheses, while t-statistics are in brackets.

*** : $p < .01$, ** : $p < .05$, * : $p < .1$

Table 9: Balance of Covariates: Scout Rankings

	Bandwidth: Rank 11-40			Bandwidth: Rank 16-35		
Variable	Rank > 25	Rank ≤ 25	Difference	Rank > 25	Rank ≤ 25	Difference
Points (per game)	9.266	10.799	-1.533***	9.567	10.261	-.696
	(4.953)	(4.777)	[-4.43]	(4.969)	(4.986)	[-1.61]
Rebounds (per game)	4.055	4.914	-.859***	4.197	4.719	-.522**
	(2.261)	(2.394)	[-5.17]	(2.226)	(2.475)	[-2.55]
Assists (per game)	1.794	1.679	.115	1.848	1.563	.286**
	(1.501)	1.339	[1.14]	(1.511)	(1.327)	[2.32]
Blocks (per game)	.541	.707	-.167***	.552	.666	-.114**
	(.646)	(.669)	[-3.55]	(.606)	(.651)	[-2.09]
Steals (per game)	.820	.907	-.0872**	.831	.865	-.034
	(.505)	(.526)	[-2.37]	(.461)	(.514)	[-.81]
Field Goals (per game)	3.315	3.853	-.538***	3.438	3.655	-.217
	(1.759)	(1.653)	[-4.44]	(1.770)	(1.711)	[-1.44]
Field Goal Percentage	.453	.472	-.019***	.454	.470	-.016**
	(.083)	(.079)	[-3.22]	(.077)	(.083)	[-2.41]
Minutes (per game)	24.981	25.759	-.779	25.576	24.537	1.039
	(8.566)	(7.572)	[-1.09]	(8.415)	(7.963)	[1.15]
N	421	371	792	274	256	530

The first and second columns for each subsample contain means with standard deviations below. The third column contains t-statistics from a t-test of unequal variances on a difference in means in the covariates. Standard errors are in parentheses, while t-statistics are in brackets.
*** : $p < .01$, ** : $p < .05$, * : $p < .1$

Table 10: Balance of Covariates: Scout Rankings, Freshmen Only

	Bandwidth: Rank 11-40			Bandwidth: Rank 16-35		
Variable	Rank > 25	Rank ≤ 25	Difference	Rank > 25	Rank ≤ 25	Difference
Points (per game)	6.707	8.616	-1.909***	7.168	7.842	-.674
	(4.066)	(4.485)	[-3.57]	(4.237)	(4.682)	[-0.98]
Rebounds (per game)	3.132	4.078	-.946***	3.393	3.722	-.329
	(1.817)	(2.165)	[-3.79]	(1.859)	(2.178)	[-1.06]
Assists (per game)	1.432	1.429	.003	1.548	1.323	.225
	(1.385)	(1.341)	[0.02]	(1.479)	(1.352)	[1.04]
Blocks (per game)	.424	.582	-.158**	.444	.520	-.076
	(.497)	(.570)	[-2.36]	(.432)	(.556)	[-0.99]
Steals (per game)	.652	.817	-.165***	.677	.768	-.091
	(.441)	(.558)	[- 2.62]	(.414)	(.557)	[-1.20]
Field Goals (per game)	2.424	3.090	-.666***	2.596	2.779	-.183
	(1.476)	(1.572)	[-3.493]	(1.515)	(1.600)	[-0.77]
Field Goal Percentage	.431	.459	-.028**	.432	.450	-.018
	(.100)	(.078)	[-2.51]	(.091)	(.079)	[-1.40]
Minutes (per game)	20.117	21.817	-1.700	21.429	19.987	1.44
	(9.269)	(8.322)	[-1.14]	(9.168)	(7.542)	[0.77]
N	128	128	256	85	85	170

The first and second columns for each subsample contain means with standard deviations below. The third column contains t-statistics from a t-test of unequal variances on a difference in means in the covariates. Standard errors are in parentheses, while t-statistics are in brackets.
*** : $p < .01$, ** : $p < .05$, * : $p < .1$

Table 11: Regression Probability Jump and Kink Estimates, Rivals Sample

Variable	Freshmen Only		Full Sample	
	11 - 40	16 -35	11 - 40	16 - 35
Five Star? (jump)	-0.021	0.128	-0.104	0.038
	(0.061)	(0.093)	(0.070)	(0.106)
Five Star × (Ranking - 25) (kink)	-0.028***	-0.038***	-0.016***	-0.027**
	(0.006)	(0.014)	(0.006)	(0.012)
Constant	0.002	-0.117	0.189***	0.071
	(0.027)	(0.071)	(0.050)	(0.080)
Wald $\chi^2(3)$	24.92	10.17	33.84	6.32
N	280	190	877	610
R-squared	.117	.037	.048	.012

This table produces second-stage estimates of (7) for players ranked within the Rivals database. All estimates are local linear and are produced for two different bandwidths, 11-40 and 16-35. Standard errors are clustered on the difference between the ranking and the threshold.
$^{***}: p < .01$, $^{**}: p < .05$, $^{*}: p < .1$

Table 12: Regression Probability Jump and Kink Estimates, Scout Sample

Variable	Freshmen Only		Full Sample	
	11 - 40	16 -35	11 - 40	16 - 35
Five Star? (jump)	-0.185**	-0.275**	-0.117	-0.209*
	(0.087)	(0.128)	(0.082)	(0.120)
Five Star × (Ranking - 25) (kink)	-0.013*	-0.007	-0.003	0.014
	(0.008)	(0.013)	(0.006)	(0.011)
Constant	0.140**	0.196**	0.219***	0.304***
	(0.066)	(0.094)	(0.052)	(0.077)
Wald $\chi^2(3)$	24.92	10.17	33.84	6.32
N	253	168	787	527
R-squared	.078	.032	.032	.010

This table produces second-stage estimates of (7) for players ranked within the Scout database. All estimates are local linear and are produced for two different bandwidths, 11-40 and 16-35. Standard errors are clustered on the difference between the ranking and the threshold.
$^{***}: p < .01$, $^{**}: p < .05$, $^{*}: p < .1$