unlevel playing fields

institutional inequality in college basketball

brian powell

February 2020

McGill

.txtLAB COLLABORATIONS

Introduction

Inspired by investigations into institutional inequality in academic hiring and publishing, this report examines the distribution of colleges attended by NBA draftees to **determine if certain institutions are disproportionately more likely to produce successful players**. I explore two types of questions:

> 1. *inequality*
>
> How do athletics fare with respect to academics in terms of institutional concentration?
>
> 2. *quality*
>
> Does attending a "top school" translate into higher performance at the professional level?

Two economic metrics - the Gini coefficient and the CR10 - are used to quantify inequality, while NBA performance is measured using the advanced stat called Value Over Replacement Player (VORP).

Overall, my findings suggest that athletics are significantly less institutionally concentrated than academics, but that **athletic concentration has undergone an important shift upwards since 2006**. I hypothesize that this is due to a rule change in the NBA prohibiting high school students from directly entering the draft. Thus, an inquiry into the fairness of college athletics has led to an insight into an unintended consequence of a rule designed to protect younger players. The effect is a growing concentration of power among a fewer number of elite schools.

Considering the financial benefits of boasting a top basketball program and the many controversies afflicting the NCAA and its constituents, this rise in inequality may provide evidence that college basketball offers yet another example of the rich-getting-richer world we increasingly inhabit. Finally, despite this rise in concentration, I show that top schools contribute only modest effects to overall player performance.

Data Collection

All data was collected from Basketball-Reference.com, which has a database cataloguing every NBA draft since its inception in 1947, including draft order and each player's career statistics[1]. I used the 1989 draft as a starting point since it was the first to introduce the 2-round system still in place today. Note that 2018 is the last year included in the data set.

To account for relocation and rebranding, I updated obsolete franchise names and locales with their current equivalent, moving the Nets to Brooklyn and swapping Bobcats for Hornets. The resulting data set features 1,742 players from 30 franchises, 27 of which participated in the 1989 draft, with an additional 3 teams introduced over the next 15 years.

Note that 309 draftees (18%) did not attend an American university. These athletes either entered the draft directly from high school – a practice that has been prohibited since 2006, but more on that later – or played outside the US before making the jump to the NBA.

Measures of Inequality

Gini Coefficient

Economists use the Gini coefficient to measure income inequality. It assesses the degree to which a given population's income or wealth distribution deviates from an ideal 1-to-1 relationship, where each person earns the same amount. Values range between 0 and 1, with **0 representing perfect equality**, since it corresponds to the case where there is no dispersion between the ideal and the actual distributions, and **1 representing perfect inequality** where one person holds all the wealth. By way of example, in 2013, the Gini coefficient was .34 for Canada and .41 for the U.S.

Here I model "income" as draft picks and "individuals" as schools. The more draft picks a school accrues the "wealthier" it can be thought to be in terms of institutional prestige. Table A1 in the Appendix lists the top 20 schools in terms of the number of draft picks they accrued between 1989 and 2018. In order to compare the NBA draft to academic publishing, I will focus on the years 1989 to 2015, where the two data sets overlap.

The Gini coefficient allows us to see how much imbalance there is among schools who have produced at least one draft pick since 1989:

0.594

The Gini coefficient for the 1,562 NBA draft picks between 1989 and 2015

If the NBA draft was a country, it would place fourth highest in terms of inequality (among countries for which estimates are available).

How does academic publishing measure up?

In this case, once again we treat each "individual" as an institution, but now "income" is defined as the number of articles published by its PhD graduates. Here we take into account 3,593 articles written by 2,439 total authors from 282 PhD-granting institutions[2]:

0.785

The Gini coefficient for the academic articles in top humanities journals between 1989 and 2015

Academic publishing has a Gini coefficient 25% higher than South Africa, the most unequal country on earth.

Thus, while the disparity between colleges in the NBA draft is comparable to some of the most inequitable countries in the world, the distribution of academic publications is even more skewed towards high-prestige schools.

Lorenz Curves

The following Lorenz curves (Fig. 1) illustrate the inequality in both domains by providing the fraction of "individuals" required to attain a certain percentage of the total "income." For example, **the top 5% of NCAA programs produced over a quarter of all draft picks** from 1989 to 2015. On the other hand, the **same proportion of PhD-granting institutions accounted for a staggering 58% of all articles published**. Thus, when we compare athletics with academics, we see a considerably higher level of institutional inequality in the academic realm. Universities are significantly more equally represented among players drafted into the NBA than they are in terms of articles published in elite journals within the humanities.

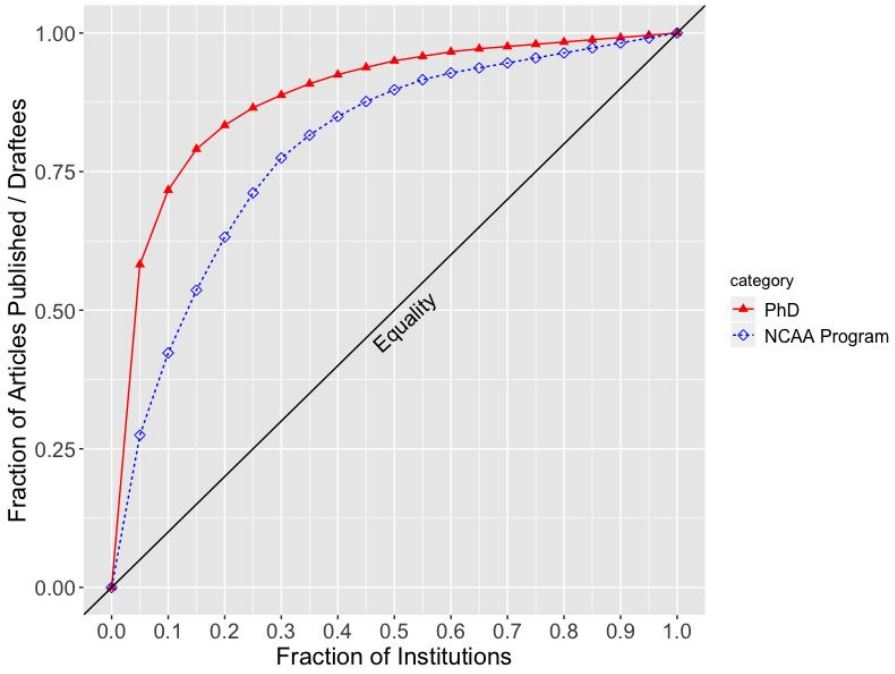


Figure 1: Lorenz curves for academic publishing and the NBA draft between 1989 and 2015

What about schools with no draft picks?

In the above models, I only consider schools that have produced a draft pick or authors who have published an article in elite journals. This does not however account for all possible schools that *could have* produced draft picks (or publishable articles). If we do this for the draft data, and divide the data around the year 2006 when the NBA instituted its rule change disallowing players to be drafted out of high school, we see a noticeable shift upward in the Gini coefficient for college basketball (Table 1). What this tells us is that the number of schools producing draft picks appears to be shrinking. The overall level of inequality among institutions has remained the same, but **the degree of concentration among a few institutions appears to be growing**.

Period	Gini Coefficient excluding Schools with 0 Picks	Gini Coefficient including Schools with 0 Picks
1989-2005	0.525	0.721
2006-2018	0.526	0.815

Table 1: Gini coefficients for select time periods. In the right column, I include all 351 schools with a Division 1 basketball program.

As a way of getting at these underlying differences, I use a measure known as a **concentration ratio**, which is used by economists to understand how much concentration there is in a given industry. What proportion of "wealth" (here draftees) is accounted for by a fixed number of institutions? A CR10 for instance calculates the **proportion of observations (i.e. draftees) accounted for by the top 10 firms (i.e. colleges)**. Note that the contents of the top 10 (the list of schools) is based on a particular time period, and so the drafts between 1989 and 1993 yield a much different selection of schools than those between 2014 and 2018. If we look at a five-year rolling average for CR10, we see the following trend (Fig. 2).

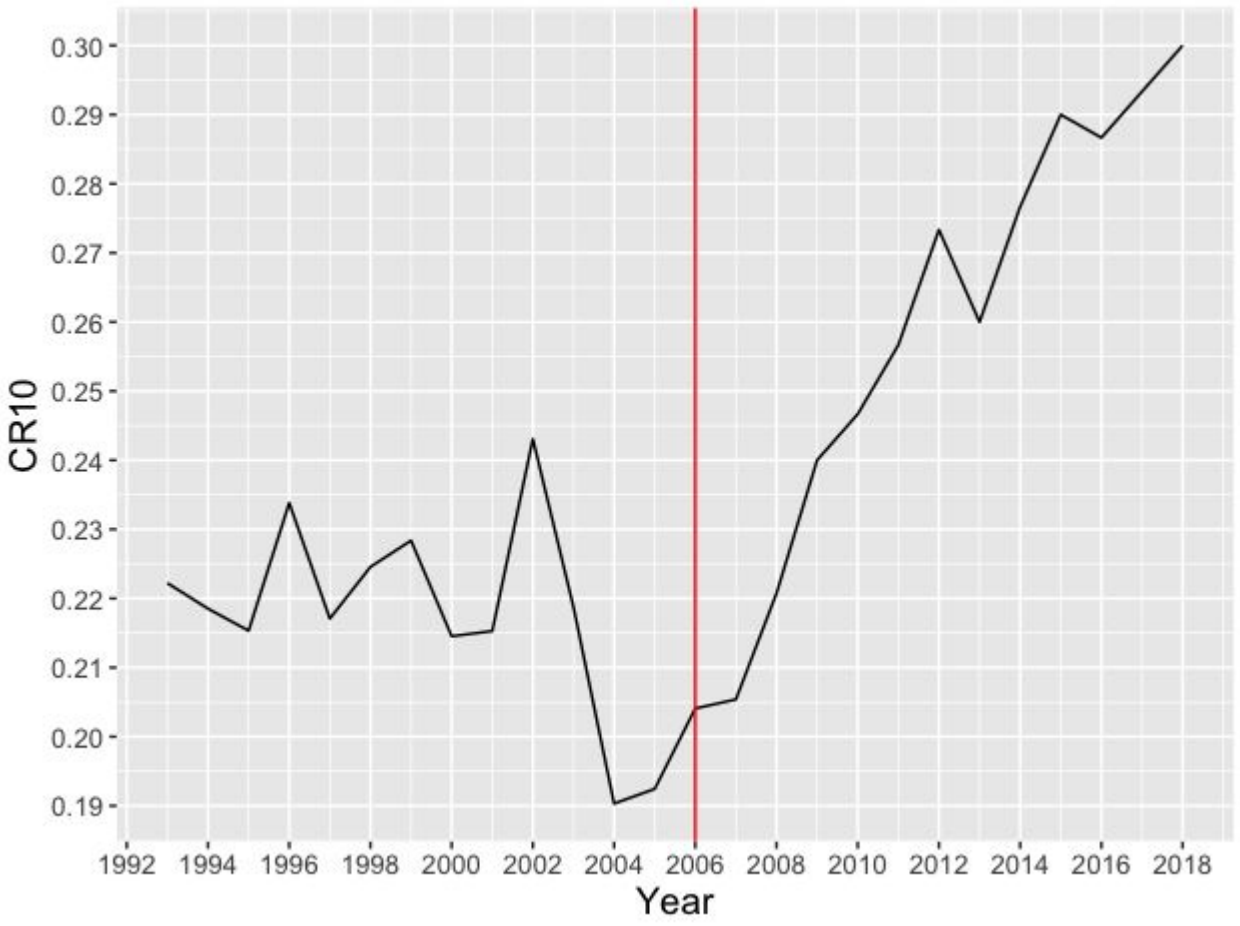


Figure 2: 5-year rolling CR10 averages for colleges in the NBA draft

The One-and-Done Rule

As previously mentioned, from the 2006 draft onward, the NBA has required players to wait a year after their high school graduation before declaring themselves eligible for selection. Athletes that could have previously entered the draft after their senior year are now obliged to spend 12 months in limbo, either attending university or playing overseas. Some colleges, most notably the University of Kentucky and Duke, now recruit highly touted prospects with the promise of providing the best platform to showcase their talents for a single year, before they jump ship for the draft. My results suggest that **this so-called "one-and-done rule" has transformed the NBA draft process in significant ways**.

For example, I examined the schools attended by draftees in five-year windows. Figure 3 charts the evolution of the number of colleges with a draftee, as well as the number of draftees from the top 10 programs, where the x-value represents the last year in that particular period. For instance, between 1989 and 1993, the top 10 schools (in that time frame) produced 60 draftees.

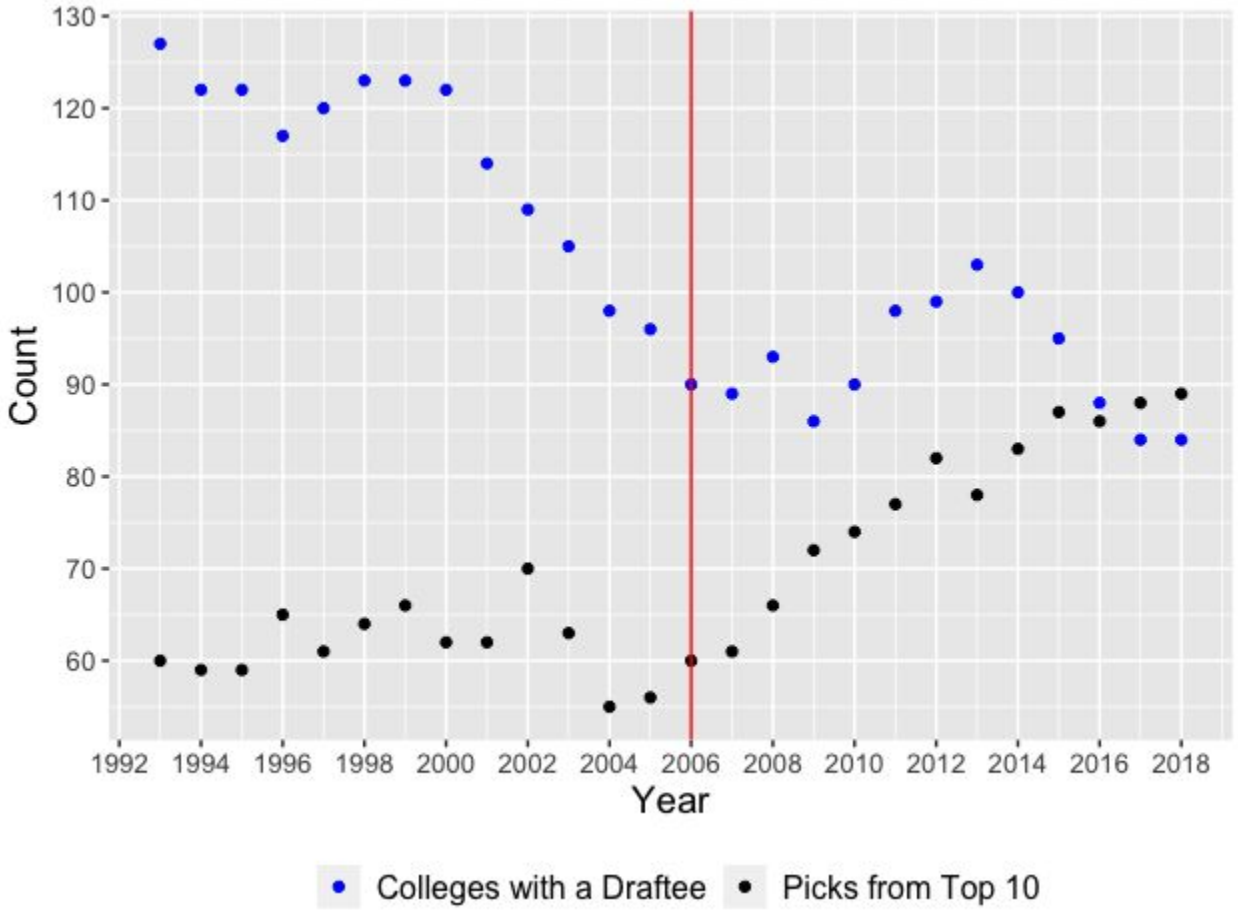


Figure 3: 5-year rolling counts for colleges with a draftee, as well as picks from the top 10 schools during that period

The results are fairly striking: the number of colleges with at least one draft pick has decreased from almost 130 from '89 to '93 to less than 85 in the past five drafts[3]. On the other hand, the top 10 schools from 2014 to 2018 now produce almost 50% more draftees than the top 10 from 1989 to 1993. This second measure was fairly constant until 2006, after which it has increased significantly.

2013-17 was the **first** 5-year period when the number of draftees from the top 10 schools **outnumbered** the number of colleges that had a player drafted

Do Top Programs Have an Effect on Player Performance?

So far, I have examined inequality with respect to the distribution of colleges in the NBA draft. However, do these top schools produce players with more successful professional careers? In other words, is there statistical evidence that draftees from Duke, Kentucky, UCLA, and other "blue blood" programs perform better at the NBA level?

To answer these questions, I evaluate the quality of players produced by a given university using each draftee's **Value over Replacement Player (VORP)**. VORP is an advanced statistic established by Daniel Myers and featured at Basketball-Reference.com. It measures the value a given player provides compared to the average bench replacement. VORP takes into account 2 factors:

1. **Box Plus/Minus (BPM)**:
This metric measures the amount of points a player contributes compared to the league average per 100 possessions. Myers provided a helpful rule of thumb: "0.0 is league average, +5 means the player is 5 points better than an average player over 100 possessions (which is about All-NBA level), -2 is replacement level, and -5 is really bad."

2. **% of Minutes Played:**
A player's VORP also considers the percentage of minutes he played out of the team's total to ensure that players with high VORP values are consistent contributors to their team's success.

Does VORP pass the eye test?

Rank	Player	Career VORP
1	LeBron James	124.9
2	Kevin Garnett	94.0
3	Tim Duncan	89.3
4	Jason Kidd	78.2
5	Chris Paul	75.9
6	Shaquille O'Neal	74.0
7	Kobe Bryant	72.1
8	Dirk Nowitzki	67.1
9	Gary Payton	63.0
10	Paul Pierce	61.5

Table 3: Top 10 career VORP values for players drafted since 1989 (through the 2017-18 season)

While there is much to be debated here, VORP clearly acknowledges LeBron James's dominance, and also rounds out its top 10 with current or future hall-of-famers.

ASIDE*: *Historic Seasons by VORP

Rank	Player	VORP	Season
1	Russell Westbrook	12.42	16-17
2	Michael Jordan	11.98	88-89
3	Michael Jordan	11.81	87-88
4	LeBron James	11.57	08-09

Quantifying Early-Career Performance

Starting with the assumption that a college's "effect" does not last throughout a player's career, I use a **3-year VORP** to measure draftees' performance, which I define as their total VORP in their first 3 seasons within 4 years of their draft[4]. Thus, players are not penalized for missing a season due to injury or for spending a year honing their skills without making an impact on the court. For example, Oklahoma star and 2009 first overall pick Blake Griffin sat out the 2009-10 campaign with a stress fracture in his knee. Fortunately, this setback did not hinder his development, as he played all 82 games and won Rookie of the Year the following season.

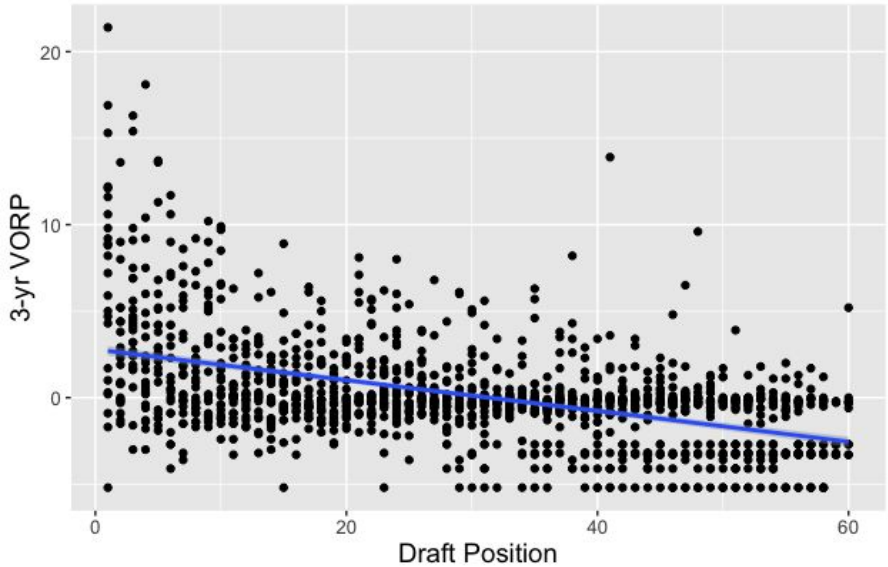
Using a box-score-based stat inherently assumes that a draft pick actually played an NBA game, within 4 years of being drafted, and so the next obstacle is determining what 3-year VORP value should be assigned to the **207 draftees with no NBA appearances** during that period. I will proceed on the following premise: 0-game draftees were deemed inferior to all players drafted within two years of their class, and so their 3 year VORP should be the minimum 3-year VORP of any draftee in that 5-year window[5]. For example, Tiny Gallon, Milwaukee's second-round pick in 2010 out of the University of Oklahoma, has never graced an NBA court. As such, he was assigned a 3-year VORP of -2.7, the minimum value for any player drafted between 2008 and 2012.

To accommodate for this 5-year window for 0-game players and the four seasons used to calculate 3-year VORP, and so I had the drafts between 1991 and 2014 at my disposal. This set of **1,394 players** provided the basis for the following insights.

VORP and Draft Position

First, I explore how 3-year VORP correlates with draft position. Namely, **do top picks outperform their peers from the start of their NBA careers?**

The following graph (Fig. 4) plots 3-year VORP by draft position, including a regression line to represent their relationship using a linear model.



This regression line has a slope of -0.089 ($t = -20.61$, $p < 2.2 \times 10^{-16}$).

This means that a **1st overall pick has a 3-year VORP about 5.3 points higher than the 60th pick**, or around 1.8 points better on a yearly basis. That was about the difference between Joel Embiid (VORP of 3.3) and Nerlens Noel (1.6) in the 2018-19 season.

Figure 4: 3-yr VORP by draft position, 1991-2014

Conditioning on the best (and the mediocre)

Interestingly, conditioning on the top 10% of 3-year VORP values provides a similar distribution across draft positions as when all players are included, with a slope of -0.070 ($t = -3.337$, $p = 0.001$). Thus, **the difference in performance between high and low picks is about the same when only considering the top players**.

For the middle 50%, however, the regression line's slope of -0.001 ($t = -0.811$, $p = 0.418$) indicates that there appears to be **no association between 3-year VORP and draft position for average NBA players** (Fig. 5).

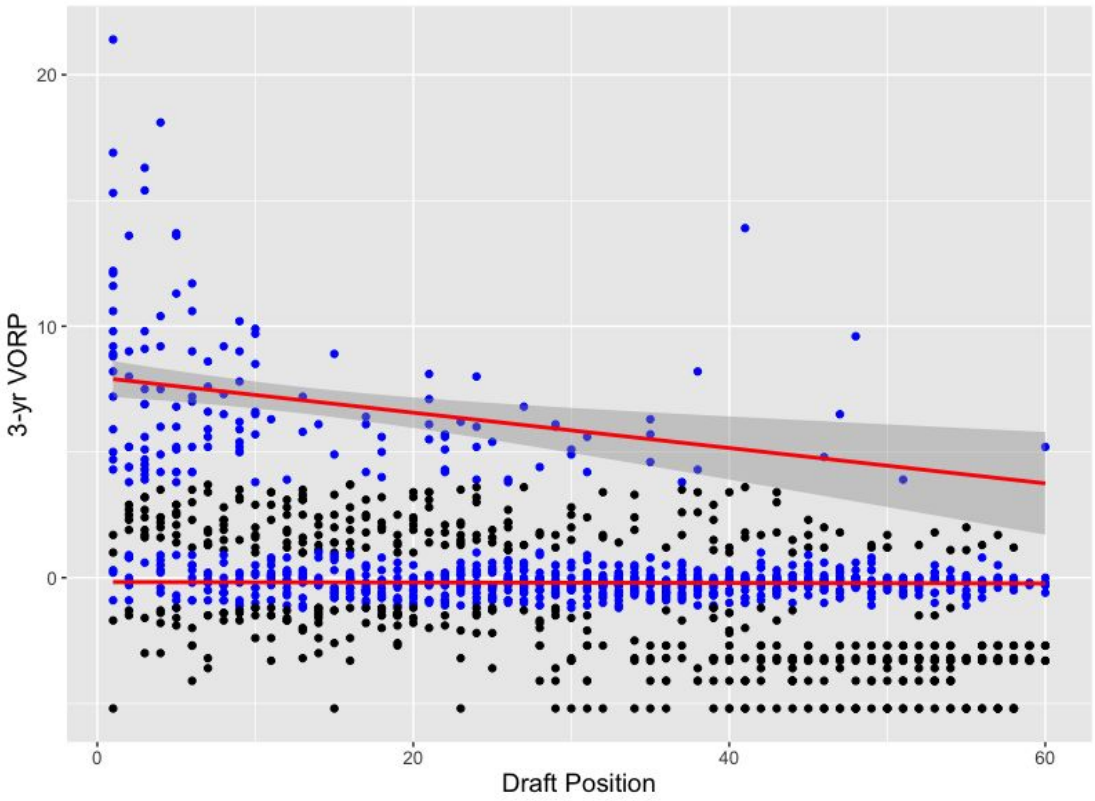


Figure 5: 3-yr VORP by draft position, top 10% and middle 50%, 1991-2014

First Round or Bust

While draft position appears to correlate with performance in general, draft *round* also has a significant impact. Given all 1,502 players drafted between 1989 and 2014, only five 1st-round picks have never played an NBA game, compared to 209 2nd-round picks, yielding an odds ratio of 57.1 (Fisher's test: $p < 2.2 \times 10^{-16}$). This provides strong evidence that **draft round is a good predictor of whether players ever play in the NBA.**

Second round picks are **57 times** more likely to never play in an NBA game than first rounders.

VORP and School Power

Next, I look at whether top colleges – as measured by their number of draftees – produce better performing NBA players.

The following graph has a point for each of the 1,138 college-attending athletes drafted between 1991 and 2014[6], with coordinates given by the number of draftees from their school's program and their personal 3-year VORP. Note that Duke and Kentucky were the top 2 colleges with 35 and 34 picks respectively.

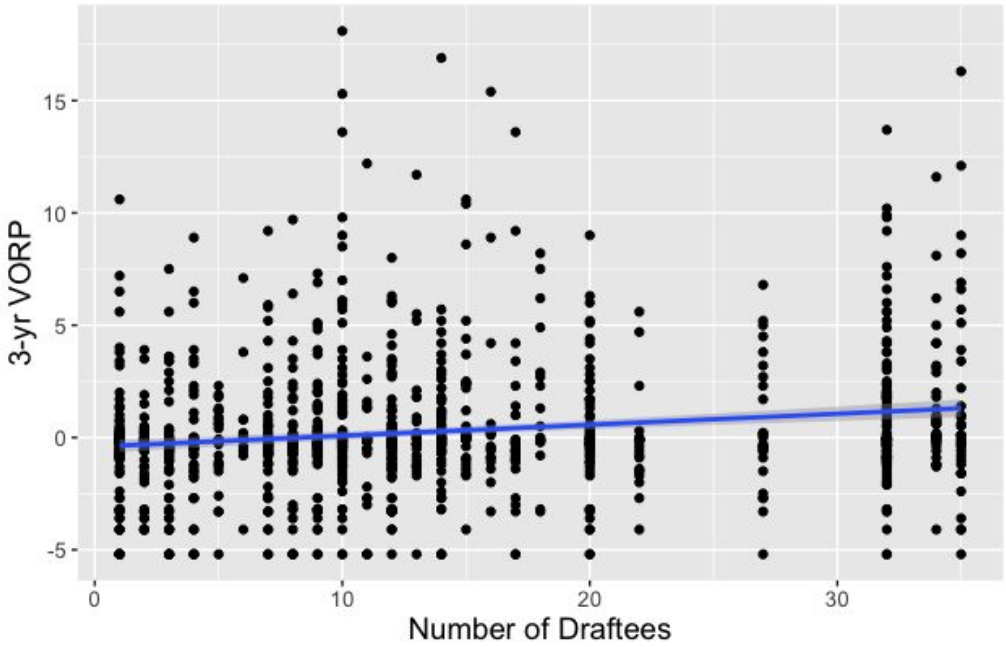


Figure 6: 3-yr VORP by number of draftees, 1991-2014

The slope of 0.049 ($t = 5.720$, $p = 1.36 \times 10^{-8}$) suggests that a player's college program has slightly less bearing on their performance than their draft position.

⇓

There is less of a difference between draftees from 1-count and 35-count schools than 1st and 60th overall draft picks.

Blue Devils vs. Cinderellas

The model above also proposes that Duke's draftees had a 3-year VORP about 1.67 points higher than an athlete that was the only pick from his school, or about 0.56 points per year. Let's put this disparity into perspective by translating back to BPM. Assuming a Duke product and a 1-count-school draftee play the same number of minutes in an 82-game NBA season, a jump of 0.56 units of VORP corresponds to a BPM 2.56 points higher. This means **the Blue Devil's team would be 2.56 points better off per 100 possessions** he plays. Considering every NBA team scored and conceded over 100 points per 100 possessions last season, this contribution is not a drastic increase.

Are top schools penalized for having more draft picks?

Since the top 10% of players account for most of the discrepancy in VORP by draft pick, a college with multiple draftees in the same year could be a victim of its own success. What if Duke's weaker picks are pulling the regression line downwards? When we only include the highest draft pick from a given school in each year, the slope increases slightly to 0.065. This means **the highest player drafted out of Duke should outperform a draftee from a 1-count school by about 0.74 units of VORP per year**, or 2.74 points per 100 possessions if they play the same number of minutes.

While attending a top school does translate to a slightly more productive start to an NBA career across all picks, this does not appear to be the case when only the top 10% of talent is considered.

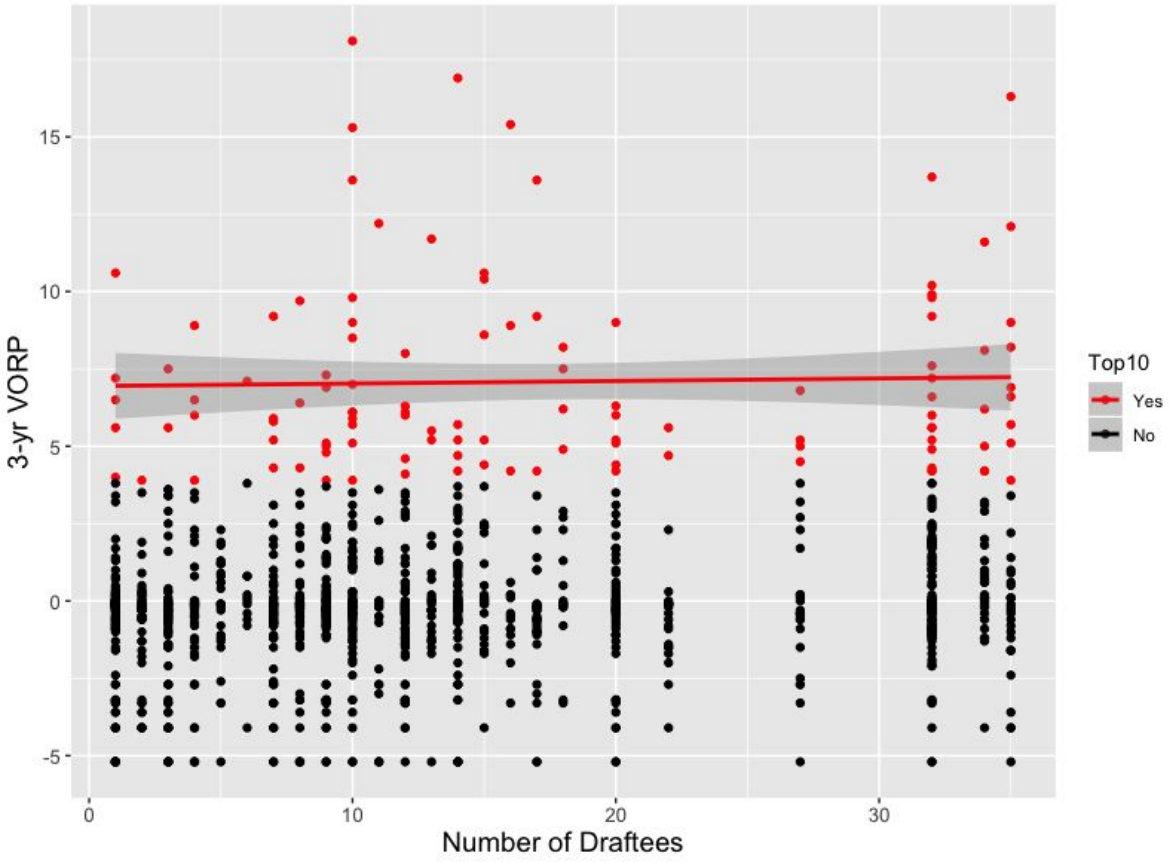


Figure 5: 3-yr VORP by number of draftees for top 10%, 1991-2014

In this case, somewhat surprisingly, the regression line has a slope of 0.008, which is close to flat.

Small programs are **about as likely** to produce top NBA players as powerhouses like Duke and Kentucky

This even distribution of talent suggests that the best high school players are not disadvantaged by attending a school with a weaker program. Furthermore, it should assuage NBA franchise's fears in taking a chance on an exceptional prospect from a lesser-known college.

Conclusion

This report has investigated the skewed distribution of colleges in the last 30 NBA drafts, particularly the rise in inequality since the introduction of the one-and-done rule in 2006. Common wisdom might suggest that athletics, with its high premium on raw physical talent, would be less meritocratic than academics. However, this turned out not to be the case. There is considerably more concentration within at least one sector of academic publishing. Nevertheless, we also see that college basketball's recent trend of the rich-getting-rich scenario mirrors larger social trends.

Since most of this paper was written prior to the 2019 NBA draft, these most recent selections provide a means of corroborating the trends observed between 1989 and 2018. In 2019, 9 of the top 10 picks played only 1 or 2 years in college - 3 of these 10 were Duke freshmen - and 5 of the top 13 were Duke or Kentucky products. In fact, 11 first-rounders attended just four schools: Duke, Kentucky, UNC, and Virginia. (It was quite the year for the ACC!)

Hopefully this attention to institutional concentration will encourage the NBA to revisit its rule change as much - if not more than - fears of yet another recruiting scandal at the college level. Creating level playing fields, even when they are hardwood courts, matters. And not just in sports.

Notes

1. NBA draft data and career statistics retrieved from Basketball-Reference.com on 9 Jul. 2018.
2. This academic data accompanied Chad Wellmon and Andrew Piper's article: "Publication, Power, and Patronage: On Inequality and Academic Publishing" (*Critical Inquiry*, 2 Oct. 2017). It can be accessed here, with caveats provided here. The four journals included in the data set are *Critical Inquiry*, *New Literary History*, *PMLA*, and *Representations*.
3. There are two possible confounding factors. First, the number of players that did not attend an American university changes from year to year, and so this fluctuation affects the total number of college draftees. Second, there were less than 60 selections in each draft between 1989 and 2004, and thus less opportunities for colleges to accrue draftees.
4. It is important to note that some players had their NBA debut more than four seasons after they were drafted, usually following a stint overseas. For example, Francisco Elson was selected 41st overall by Denver in 1999, but played 4 seasons in Spain before his first game with the Nuggets. For the purposes of this measure, these draftees are considered 0-game players.
5. Assigning the minimum within a 5-year window for 0-game draftees explains the clustering of low 3-year VORP values at the bottom of Figures 4-7.
6. 136 of these 1,138 draftees were 0-game players.

Appendix

Table A1: Top 20 Colleges for Producing NBA Draftees, 1989-2018

Rank	College	Count
1	Duke University	50
2	University of Kentucky	50
3	University of California, Los Angeles	41
4	University of Arizona	40
5	University of Kansas	39
6	University of North Carolina	38
7	University of Connecticut	30
8	Syracuse University	29
9	Michigan State University	25
10	University of Michigan	25
11	University of Texas at Austin	25
12	Georgia Institute of Technology	24
13	University of Maryland	23
14	University of Louisville	21
15	Louisiana State University	19
16	University of Florida	19
17	Florida State University	18
18	Indiana University	18
19	Ohio State University	16
20	Stanford University	16