Changes, Changes, Changes.

Vernice Lee

Southern Utah University

Why Only 4?

Abstract

Like any professional sports organization, the NBA holds a draft that allows teams to strategically choose amongst qualified individuals to participate within their organization to better them as a whole. Over the course of the history of the NBA, the system in which the draft functions has changed several times now resulting in 4 teams being chosen from a Lottery to select picks No. 1 - No. 4. The purpose of this study is to compare the wins produced by the first 4 draft picks to the following draft picks up to justify that indeed the first 4 draft picks are more beneficial than the rest. Using a simple OLS model, we will compare differences in means between the 2 groups also controlling for individual player statistics and team statistics to account for more variation in our measures of *Wins Produced.* If evidence is found that supports the idea that the first 4 picks produce more wins when compared to the picks after, there is an obvious incentive to tank or do worse to obtain a Lottery Pick.

Introduction/Literature Review

The National Basketball Association also known as the NBA has been around for more than 70 years now. At the conclusion of every season, the NBA, amongst many other professional sports organizations, hold an event known as the NBA Draft. As (Florke & Ecker 2003) noted, the NBA Draft is vital to the health of the organizations that get a chance to choose an amateur player usually coming from college, high school, in some cases, or overseas. The first ever NBA Draft was held in 1947 and since then the system of the draft and how it functions has altered several times over the history of the NBA. The way teams got the opportunity to select a player was determined from a draft order. The draft order initially ranked teams inversely according to their records at the conclusion of the exhibition season, meaning the team with the absolute worst record chose first and the team with the best record chose last.

Why Only 4?

The way the draft order is determined now differs from the methods used in 1947. In fact, the term “lottery” is now associated with the NBA Draft. The original system of the NBA Draft first altered in the 1980’s as Penrice (1995). From 1966-1984 the NBA team that finished with the worst record in each conference flipped a coin to receive the No. 1 overall pick and the other team receives the No. 2 overall pick and the remaining teams received their picks in reverse order of their regular season records. 1984 the NBA then implements to idea of the lottery (selecting at random) to determine the order of picks for the first round and the second rounds order was determined by inverse regular season records. In 1986 the Board of Governors then decided the lottery would only determine the order for the first three picks, with the rest of the teams then choosing inversely according to their regular season records. This would assure that the team with the worst record would choose no later than the 4th pick. As the years continue, the NBA changes the draft and its system about 5 more times to bring us to the current rules that are in place today.

Now the way the lottery is set up today involves the chances of receiving the No. 1 overall pick to be weighted depending on your regular season record. The teams with the three worst regular season records will each have a 14 percent chance of winning the lottery. The following 4th seed will have a 12.5 percent chance, 5th seed will have a 10.5 percent chance, and 6th seed will have 9 percent chance. After the first four picks are received, the rest of the teams choose according to the inverse order of their regular season records.

The exact process for the actual draft is noted as:

“Fourteen ping-pong balls numbered 1 through 14 will be placed in a lottery machine. There are 1,001 possible combinations when four balls are drawn out of 14, without regard to their order of selection. Before the lottery, 1,000 of those 1,001 combinations will be assigned to

the 14 participating lottery teams. The lottery machine is manufactured by the Smart Play Company, a leading manufacturer of state lottery machines throughout the United States. Smart Play also weighs, measures and certifies the ping-pong balls before the drawing.
The drawing process occurs in the following manner: All 14 balls are placed in the lottery machine and they are mixed for 20 seconds, and then the first ball is removed. The remaining balls are mixed in the lottery machine for another 10 seconds, and then the second ball is drawn. There is a 10-second mix, and then the third ball is drawn. There is a 10-second mix, and then the fourth ball is drawn. The team that has been assigned that combination will receive the No. 1 pick. The same process is repeated with the same ping-pong balls and lottery machine for the second through fourth picks. If the same team comes up more than once, the result is discarded and another four-ball combination is selected. Also, if the one unassigned combination is drawn, the result is discarded and the balls are drawn again. The length of time the balls are mixed is monitored by a timekeeper who faces away from the machine and signals the machine operator after the appropriate amount of time has elapsed.", found directly from NBA.com.

Extensive research has been conducted to analyze the NBA player performance, production of wins, the NBA Lottery, and even the idea of "tanking" to potentially receive a better draft order position by purposefully losing games. Interesting results have been discovered and some of these findings are the foundations in which future research will build off of.

The first case in which teams were identified as facing incentives to lose games and acting upon those incentives was discovered Taylor and Trogdon (2002). Further research has been discovered by Price, Soebbing, Berri and Humphreys (2010) which does indeed find evidence that NBA teams were more likely to intentionally lose games at the conclusion of the regular season during seasons where incentives were the largest. Empirical analysis using simple

linear regressions also found evidence of incentives to tank done in a similar study done by Walters and Williams (2012). An interesting finding using simple fixed-effects methodologies indicates “Playoff-eliminated teams are 14 percentage points more likely to lose a game when doing so increases their chances in the draft lottery.” (Walters and Williams, 2012). Another incentive to tank that was noted was the idea of increasing one's attendance the following season which seemingly comes from the publicity and anticipation of the first draft pick. Being an early or first draft pick leads to the idea of performing better than everyone else available to be drafted.

Measuring a basketball player’s performance and determining their true value and contribution has been a discussion of concern for years. A simple study done by Mikolajec, Maszczyk, and Zajac (2013) will note that a few key performance indicators are things such as *3rd quarter PPG, Win %, Offensive EFF, Avg Fouls, Avg Steals,* and so on.

Combining the ideas and lessons learned from the presented studies, I will try to determine what are the effects of being a lottery pick. My contribution to this topic of research will implement important findings regarding individual player performance measures and challenge the system of the lottery to ensure it’s intended purpose is actual something real.

Data

The NBA collects and reports statistics on every individual player there is to come through the league. These statistics are readily available to the public which allows people not associated with the NBA to conduct research ourselves. Important research that lays the foundation for the topic of interest in this study comes from *The Wages of Wins* by David J. Berri in 2010. The data that will be used for this paper comes from Berri (2010) compiled from

Why Only 4?

NBA.com. The dataset has since been updated to include draft data which dates back to the 1973-1974 season and now through the 2017-2018 season.

The dataset consists of a numerous number of variables that will be considered and controlled for in order to explain the topic at interest. Although the statistics date back to the 1973-1974 season, an area of interest is the 1984 season which is the first season the lottery was implemented. A later change to the draft will be made in 2017 but only the weights for some of the teams have altered. Some of the variables that will be considered and measured can be grouped into two different categories. One category of variables will concern individual player performance like *Age, Steals, Assists, Turnovers, Total Rebounds,* and so on. The other category being team performance will include variables such as *Teams Wins, Team Field Goals Made, Team Assists, Team Steals,* and many more.

Summary Statistics

	No. 1 Mean	No. 2 Mean	No. 3 Mean	No. 4 Mean	No. 5 Mean	No. 6 Mean	No. 7 Mean	No. 8 Mean
Wins Produced	**6.72**	**5.152977**	**4.985884**	**4.412279**	**5.31863**	**2.340731**	**3.426171**	**2.943514**
Age	**26.79**	**26.75074**	**26.60119**	**26.78962**	**27.32698**	**25.73622**	**26.05363**	**26.35714**
Games	**61.69**	**65.24036**	**62.94048**	**63.39071**	**65.43324**	**60.94882**	**64.14826**	**62.22789**
Games Started	**54.39**	**48.01484**	**48.11446**	**42.94536**	**47.20163**	**36.2126**	**42.19558**	**35.19388**
Minutes played	**2032.91**	**1905.585**	**1937.807**	**1859.325**	**1927.534**	**1587.669**	**1773.744**	**1621.483**
Field Goals made	**421.67**	**323.0356**	**361.2232**	**316.9809**	**342.4223**	**245**	**291.735**	**255.8231**
Field Goals attempted	**851.91**	**690.9466**	**780.6875**	**697.9344**	**743.7384**	**543.6496**	**646.7161**	**571.7381**
Offensive Rebounds	**139.27**	**108.3947**	**84.73512**	**96.37705**	**97.32425**	**78.73228**	**76.8959**	**81.21429**
Defensive Rebounds	**356.55**	**284.8991**	**245.6786**	**253.0765**	**262.4142**	**193.9961**	**205.3691**	**194.3639**
Assist	**186.49**	**183.6439**	**216.7381**	**182.306**	**193.624**	**125.4016**	**190.1735**	**144.1871**
Blocks	**84.86**	**59.81899**	**41.58036**	**41.71038**	**37.13624**	**32.87008**	**30.75394**	**32.63605**
Turnovers	**149.49**	**121.3769**	**136.397**	**115.623**	**127.8529**	**92.64173**	**115.8644**	**98.08163**
Points	**1104.51**	**846.7656**	**987.3423**	**845.6011**	**923.8937**	**650.8425**	**780.3407**	**678.3503**

Why Only 4?

The statistics shown above demonstrates that there is variation amongst the data, thus satisfying one of our assumptions when determining if our estimates will be biased. We indeed have data that varies across the board for most variables, have someone who played a maximum number of games of 87, as well as having observations that only played 1 or 2 games. Another assumption that is satisfied for proof of unbiased estimates is the fact that the dataset contains ever single observation there is in the NBA, which means random assignment isn't an issue when wondering if we have a representative sample.

The dependent variable of interest to explain our concerns of the lottery happens to be *Wins Produced*. The way *Wins Produced* is computed and measured is explained in *Stumbling on Wins: Two Economists Explore the Pitfalls on the Road to Victory in Professional Sports.* The measurement stems from Chris Paul (2008-09), but then adjust for production of teammates and position player to calculate a value for *Wins Produced.* The measure will allow how us to evaluate the productivity of each player in each season they participate in. This variable is important and allows us to analyze and measure how many wins each individual draft pick will produce for his team. Wins are the obvious goal for every single team in the NBA so the measure of how many wins a single player produces can be vital when analyzing a team or player and making important decisions regarding the organization. In addition to the summary statistics above, T-Test were ran across the *Wins Produced* means for each draft pick to see if they are statistically different and indeed a few of them are. This simply means that the average number of wins produced by the No. 1 draft pick is statistically different from the average number of wins produced by the No. 2 draft pick.

The important independent variable of interest that will be used to answer our question will be a dummy variable for *Lottery Pick* which will be generated from *Draft Position. Lottery*

Why Only 4?

Pick will equal 1 if the individual player was drafted No. 4 overall or earlier (*Draft Position <= 4*) and will equal 0 if the player was drafted after the 4th pick but no later than the 8th pick. The reason for this cut off is because the sample I want to look at are the teams with at least a 5% chance of winning the lottery picks, with the 8th worst regular season record team having 6% chance. This will compare the lottery pick teams to the teams who lost the opportunity to be a lottery pick.

One possible issue that will be addressed is some players didn't play a single game the following season after being drafted and it's important we understand and note that that. An example of this case is Ben Simmons, who was the No. 1 draft pick in the 2016 NBA Draft who suffered a fractured foot and did not compete in a regular season game. In that season, he obviously produced 0 wins which is why we will look at the *Wins Produced* for each individual season as an observation to see what effect these players will have each season.

Method

With the dataset obtained, the ideal tool that needs to be utilized is regression analysis, which will provide useful calculations to better understand areas of interest. The following model will be the foundation that will be built upon to analyze and consider other areas of concern we have when looking at individual wins produced.

$$(1) \qquad Wins\ Produced = \beta_0 + \beta_1(Draft\ Position) + u$$

Using this univariate OLS model, β_1 will give us an estimate of the change in wins produced for someone who is drafted within the first 4 picks compared to the following picks up until the 8th pick which is where the cutoff is. There will be factors to consider that play a role in both *Wins Produced* and *Draft Position* that will try to be accounted for. Accounting for individual player statistics to help measure performance and team statistics to measure the team's

overall performance, and a measure of both, will allow us to try and minimize the factors that may be continued in our error term. We will run a handful of regressions to analyze the effects that come from the additional control variables.

(2) $Wins\ Produced = \beta_0 + \beta_1(Draft\ Position) + \beta^*(Player\ Statistics) + u$

(3) $Wins\ Produced = \beta_0 + \beta_1(Draft\ Position) + \beta^*(Team\ Statistics) + u$

(4) $Wins\ Produced = \beta_0 + \beta_1(Draft\ Position) + \beta^*(Team\ Statistics) + \beta^*(Player\ Statistics) + u$

The first regression, Figure (1), will simple regress *Wins Produced in 48 Minutes* on *Draft Position* as a continuous variable. Next, we will add just individual player statistics to the regression, Figure (2), followed by just team statistics, Figure (3), and finally a combination of both individual player statistics and team statistics, Figure (4). This will allow us to build on and adjust our independent variable of interest to a dummy variable, *Lottery Pick*. The value of *Lottery Pick* will be 1 if the *Draft Position* is less than or equal to 4, and a 0 if the pick is anything greater than 4 but also less than the 9th pick. The will allow us to determine if indeed choosing within the first 4 draft picks will produce more wins when compared to the picks that follow up until the 8th pick.

(5) $Wins\ Produced = \beta_0 + \beta_1(Lottery\ Pick) + \beta^*(Player\ Statistics) + u$

(6) $Wins\ Produced = \beta_0 + \beta_1(Lottery\ Pick) + \beta^*(Team\ Statistics) + u$

(7) $Wins\ Produced = \beta_0 + \beta_1(Lottery\ Pick) + \beta^*(Team\ Statistics) + \beta^*(Player\ Statistics) + u$

Results

Table 3. Regression Results for *Draft Pick*

VARIABLES	(1) Wins Produced	(2) Wins Produced	(3) Wins Produced	(4) Wins Produced
Draft Pick	-0.493***	0.0200	-0.406***	0.0222
	[0.0469]	[0.0168]	[0.0451]	[0.0159]
Control for Individual Player Statistics	No	Yes	No	Yes
Control for Team Statistics	No	No	Yes	Yes
Constant	6.682***	2.364***	3.704***	4.123***
	[0.230]	[0.316]	[1.155]	[0.469]
Observations	2,628	2,628	2,628	2,628
R-squared	0.040	0.901	0.136	0.914

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 3. displays the results of the first set regressions ran while considering *Draft Pick* as the important independent variable of interest. In Column (2) we control for individual player statistics, in Column (3) we control for team statistics, and lastly in Column (4) we control for both team and individual player statistics. When controlling for individual player statistics, we can see that the R^2 of our model jumps up to 90% which means that the individual player statistics helps explain about 90% of the variation in *Wins Produced.* Although the coefficient for *Draft Pick* isn't statistically significant when controlling for individual player statistics, the direction of the coefficient is positive which is interesting. This would imply that as you increase your draft pick, the players will produce more wins on average. One would expect the draft picks that are drafted earlier will produce more wins than that the ones to follow. The next set of regressions will help explain the impact of being a lottery pick in particular on wins produced.

Table 4. Regression Results for *Lottery Pick*

VARIABLES	(1) Wins Produced	(2) Wins Produced	(3) Wins Produced	(4) Wins Produced
Lottery Pick	1.668***	-0.121	1.350***	-0.151**
	[0.215]	[0.0735]	[0.206]	[0.0694]
Control for Individual Player Statistics	No	Yes	No	Yes
Control for Team Statistics	No	No	Yes	Yes
Constant	3.651***	2.525***	1.410	4.297***
	[0.157]	[0.304]	[1.155]	[0.463]
Observations	2,628	2,628	2,628	2,628
R-squared	0.022	0.901	0.123	0.914

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 4. demonstrates the results from the from the second set of regressions ran while looking at *Lottery Pick* as the important independent variable of interest. As explained in the results from Table 3. The R^2 of the model increases to about 90% when controlling for individual player statistics. Column (4) of Table 4. actually, displays an interesting finding. When controlling for individual player statistics and team statistics, the coefficient for *Lottery Pick* actually becomes statistically significant at the 5% level with a T-Stat of 2.18 and a P-Value of 0.0029. This means that on average, being a lottery pick will produce about .15 wins when compared to not being a lottery pick.

Conclusion

Since the system of the NBA Lottery Draft has been implemented, there has always been speculation and concern. Simple changes have been made over the years, and ideally the changes are for the better but our findings lead to a different result. When controlling for individual player statistics and team statistics, the coefficient for *Lottery Pick* is -0.1511 and is also statistically significant at the 5% level. This implies that on average being a lottery pick, picked within the first 4 picks, produces less wins when compared to someone who isn't a lottery pick.

There is statistical evidence that supports the idea that then system of the NBA draft as of right now aren't producing wins. Teams were found to theoretically tank in previous years to receive a higher draft pick, which should help produce wins. Future studies may build off of this by figuring out what to arguing for the reason why players who aren't lottery picks produce more wins. Also, we assume that coaches and organizations make rational and smart decisions when choosing these draft picks, but there is no way to determine and measure those. To conclude this study, through empirical analysis, we prove that being a lottery pick produces less wins, which essentially means under the new system of the NBA Lottery Draft, there is no incentive to tank anymore.

References

Berri, D. J. (1999). Who is 'most valuable'? Measuring the player's production of wins in the National Basketball Association. *Managerial and decision economics*, *20*(8), 411-427.

Florke, C. R., & Ecker, M. D. (2003). Nba draft lottery probabilities. *American Journal of Undergraduate Research*, *2*(3), 19-29.

Mikołajec, K., Maszczyk, A., & Zając, T. (2013). Game indicators determining sports performance in the NBA. *Journal of human kinetics*, *37*(1), 145-151.

Price, J., Soebbing, B. P., Berri, D., & Humphreys, B. R. (2010). Tournament incentives, league policy, and NBA team performance revisited. *Journal of Sports Economics*, *11*(2), 117-135.

Penrice, S. G. (1995). Applying elementary probability theory to the NBA draft lottery. *SIAM review*, *37*(4), 598-602.

Walters, C., & Williams, T. (2012). To tank or not to tank? Evidence from the NBA. In *MIT Sloan Sports Conference March. Boston, MA*.