

Contract Year Effect in the NBA*

Changhao Shi[†] Jonathan Liu[‡] Terry II Culpepper[§] Sean Choi[¶]

June 10, 2020

Abstract

We investigate whether being in a contract year has an effect on NBA players by motivating them to put in more effort on playing. We use player statistics such as distance covered on the field per minute as a proxy for effort. We tackle three main challenges arising from the data: lack of a good proxy for effort, inter-correlation of effort between players, and player-specific effects on player statistics and outcomes. Using an OLS regression controlling for individual fixed effects and a double LASSO regression, we found that being in a contract year has no statistically significant impact on various measures of player effort.

Keywords: Sports Economics, Athletes, Contracts

JEL Codes: Z20, Z22

*We would like to thank Dr. Hortaçsu and Francisco del Villar Ortiz Mena for their unique insights.

[†]University of Chicago

[‡]University of Chicago

[§]University of Chicago

[¶]University of Chicago

1 Introduction

2 Literature Review

3 Data

The NBA salaries of every player from the 2016-2017 season to the 2019-2020 season were collected from [basketballreference.com](https://www.basketballreference.com). Note that [basketballreference](https://www.basketballreference.com) only displays data for the current season, so to get salary data for 2016-2017 to the 2018-2019 seasons, we utilized web.archive.org. Players were determined to be contract years if they did not have a salary for the following season. Furthermore, panel data of all advanced stats for each active NBA player was scraped from [basketballreference.com](https://www.basketballreference.com) for every regular season these players played in. Most of the advanced stats are measurements of a player's productivity on the court. These include the age of the player, team the player played for, position the player played for, the number of games the player played for a particular season, average minutes played per game, player efficiency rating, true shooting percentage, three-point attempt rate, free-throw attempt rate, offensive rebound percentage, defensive rebound percentage, total rebound percentage, assist percentage, free throw attempt rate, offensive rebound percentage, defensive rebound percentage, assist percentage, steal percentage, block percentage, turnover percentage, usage percentage, offensive win shares, defensive win shares, win shares, win shares per 48 minutes, offensive box plus/minus, defensive box plus/minus, box plus/minus, and value over replacement player.

Another way to measure a player's productivity/effort on the court is with boxout data. In theory, the more a player boxes out opposing players in order to grab rebounds, the more productive that player is to the team. All boxout-related data was collected from stats.nba.com for every player from the 2016-2017 to the 2019-2020 regular season. These include the number of boxouts a player averages per game, number of boxouts on offense a player averages per game, number of boxouts on defense a player averages per game, average

number of rebounds the team grabs as a result of a player boxing out, average number of rebounds the player grabs as a result of him boxing out, percentage of times a player boxes out on offensive, percentage of times a player boxes out on defense, the percentage of times the teams grabs a rebound when boxing out, and the percentage of times the player grabs a rebound when boxing out.

Another method of measuring a player's value on the court is by using data on how often the player touches the ball. Players who handle the ball more often are typically much more valuable to his team. Data for player touches was collected from stats.nba.com. These include the number of touches a player averages per game, the number of touches the player averages in the front court per game, the percentage of the time the player has the ball when he is on the court, average seconds the player has the ball when he touches it, the average number of dribbles the player takes whenever he touches the ball, the number of points the player scores per touch, the average number of times a player touches the ball in the elbow part of the court, the average number of post-ups a player has per game, the average number of times a player touches the ball in the paint, points per elbow touch, points per post-up, and points per paint touch.

Note that when analyzing the data, the datasets for player salary, advanced player stats, boxouts, and touches were merged together by player name.

3.1 title

3.2 Model and Analysis

Challenges in estimating the contract year effect fall under three categories. Firstly, player effort cannot be observed, and existing metrics such as distance covered on the field per minute may only be loosely correlated with player effort. For example, a player may put in effort by perfecting 3-point throws and so roughly run the same amount of distance but attempt more 3-point shots per game. Another player may put in effort by running longer

each game to gain tactical advantage on a field. It is also possible that the player's effort fails to translate into an observable metric; for example, he might put in effort into team-building exercises and coordinate much more on the field. This also brings us to the next challenge. The effort put in by a particular player may also correlate with other players' performance. Since basketball is a team game, effort put in by one player may synergize with effort put in by other players. We can also look at this issue game-theoretically. Consider the standard public goods game, and let y_i be the energy of the player i devoted to the match, out of a total of 1 endowed unit of energy. The payoff function for player i can be thought of as

$$\Pi_i = (1 - y_i) + \alpha \sum_j y_j$$

where $0 < \alpha < 1$. The sum of the players' effort correlates positively with the expected probability of winning a match. When $\alpha < 1$, the Nash equilibrium of this game is for all players to invest 0 energy into the match. However, when factoring in social norms, we expect players to fall into two categories. Based on empirical evidence, players will either put in more of their endowment when others put in more, or put in less when others put in more (?). Regardless, this makes estimating the contract year effect more difficult, as the effort devoted by an individual correlates with the effort of other individuals. Furthermore, this correlation is ex-ante unknown. The final challenge lies in the heterogeneity in different individuals. Even if we assume that the first two challenges are resolved, and that all players reflect their effort by pursuing, for instance, more three-point shots per minute, underlying unobserved parameters such as player skill and talent, player psychology during matches, and the correlation of the contract year effect with other covariates such as current salary means that individual fixed effects cannot be ignored.

To counteract the first challenge, we run multiple regressions with different outcome variables, such as player distance traveled per minute on the field and three point shot percentage. This partially circumvents the issue that proxies could be loosely correlated

with player effort by looking at how being on a contract year affects a wide range of player statistics, instead of an individual measurement. To counteract the second challenge, we have to assume that the underlying correlation between players due to effort is homogeneous across teams. Then we can estimate the contract year effect on teams by looking at how many members are in a contract year and look at team level statistics. Finally, to counteract the third challenge, we control for fixed effects in our regression by adding indicator variables for players or teams as a covariate.

3.3 Controlling for Fixed Effects

To control for fixed effects (omitted variables of an individual), we add indicator variables for players or teams as a covariate. This, along with the existing controls, gives us the regression model

$$y_{it} = \sum_{j=1}^N \alpha_j \mathbb{1}\{i = j\} + \mathbb{1}\{i_t \in \text{Contract}\} \beta + c'_{it} \gamma + u_{it}$$

where y_{it} is the outcome player statistics for player i in time period t , $\mathbb{1}\{i = j\}$ is the indicator variable for players, $\mathbb{1}\{i_t \in \text{Contract}\}$ is the indicator variable for whether a player is in a contract year in a given time period, and c_{it} are various controls, such as age and current salary. The identifying assumption is then that unobservable effects soaked up by the individual indicator variable that simultaneously affect the outcome player statistics and the explanatory variable and covariates are time-invariant. In other words,

$$\text{Cov}(x_{i1}, u_{it}) = \dots = \text{Cov}(x_{iT}, u_{it}) = 0,$$

where x includes both the explanatory variable and the controls. This assumption is innocuous enough in this context. We don't expect unobserved characteristics (omitted variables) of an individual to change dramatically across years that also affect whether a given player is in a contract year. For example, a player may get married and be very happy that year. This would cause him to play better, therefore increasing his performance and we will see a

change in his player statistics. However it is unlikely that it will affect whether he is in a given contract year, as that aspect largely depends on how many years ago the player signed the contract.

4 Results

5 Discussions

6 Conclusion

Tables

Figures

Appendix A. Placeholder