The Effect of NBA Contract Years on Player Performance

POLI SCI 200C Final Project

Wesley Cheng, Ami Wulf, Tianyu Ye

Abstract

The contract year phenomena has long been purported to exist across professional sports leagues in America. This project attempts to quantify the contract year effect in the NBA by using John Hollinger's player efficiency rating (PER). Our work focuses on carefully crafting the definition of a contract year player to remove many potential biases that were often ignored in previous works. Under the selection under observables assumption, we use matching and a player fixed effects model, both of which suggest that there is a statistically significant positive effect of contract years on PER. As a result, there is reason to believe that contract year performances in the NBA should be evaluated differently from non-contract year performances.

1 Introduction

NBA contracts play a crucial role in the competitive balance of the league. There is an ever increasing pressure on front offices to accurately assess player performances when constructing competitive rosters as the NBA's salary cap system puts limits on the amount teams can spend. One common phenomena that is believed to affect player performance is the contract year phenomena, which refers to players supposedly increasing their performance on the court (consciously or not) in the last year of their contract deal. The most often cited example of this supposed contract year effect is Erick Dampier's 2003-04 contract year. Up to that point, Dampier had been a serviceable starting center for the Golden State Warriors, before averaging an impressive double-double (12.3 points per game and 12.0 rebounds per game) putting up a 20.1 PER (6 points higher than his career average) during his contract year. His reward? A 7 year 73 million dollar contract with the Dallas Mavericks the following off-season. He would never come close to replicating his contract year performance for the rest of his career.

In this project, our goal is to see how being in a contract year affects NBA players' quality of play. Quantification of this effect would allow NBA teams to predict boosts in performance during contract years as well as to better assess a player's underlying, non-contract-year abilities. Throughout, we prioritize removing biases and obtaining accurate estimates rather than producing generalizable results. Our project is structured as follows. Section 2 gives a more thorough explanation of NBA contracts and their structure, our outcome of interest (PER), as well as a review of previous works on this topic. Section 3 describes the NBA data that we use in our study and clearly defines how we view the definition of contract years. Section 4 delves into our identification strategy and the methods we employ in order to estimate the ATT of being in a contract year. Our results are presented in section 5, and we make some concluding remarks including potential threats to validity and areas for further exploration in our last section.

2 Background and Previous Work

2.1 NBA Contracts

In the current day NBA, contracts typically range from one to five years. Generally speaking, the younger and more talented the player is, the longer their contracts are. A contract year refers to the last year of a player's deal. This means contract years are the last year before they hit free agency, which is a time in which they can negotiate new contracts with either their current team or other teams in the league. One can imagine that players might approach their contract year differently compared to other non-contract years due to the financial uncertainty that follows and recency bias that league front offices often exhibit when evaluating players entering free agency.

One key component in the structure of contracts is the concept of options. Contract options are somewhat analogous to stock options, in that they give either the team or player the right to extend their contract an additional year. Thus, there exists both team options and player options. Team options give the team more leverage as they give the team more financial flexibility. If the player improves their play above their current pay-grade, the team can lock in the current rate for an additional 1-2 years. Conversely, player options would allow the player to lock in their current rate even if they do not perform at the level expected of them. It is clear that in these cases, it become more difficult to evaluate whether not these players' last year should be considered a true contract year. This will be explained in more detail in section 3 when we talk about how we precisely define a contract year.

2.2 Player Efficiency Rating

Player Efficiency Rating (PER) is a all-in-one metric developed by John Hollinger to attempts to measure a player's net accomplishments and contributions over the course of a season. PER takes into account traditional statistics such as points, rebounds, assists, steals, missed shots, fouls, etc. It is adjusted for team pace and minutes played, which allows for comparisons between bench players and starters as well as players that play on teams that play with a more methodical pace (lower counting stats) compared to faster-paced teams.

As with any statistic that attempts to capture a player's performance in one single number, such as win-shares or plus/minus, PER is an imperfect measure and has its fair share of valid criticisms. One particular problem, that Hollinger himself admits, is that it is an not a reliable measure of player's defensive skills. It only takes into account steals and blocks, both of which are not perfect indicators of player's defensive performance. Take for example Andre Roberson who is considered one of the best perimeter defenders and was named to the 2016-17 NBA All-Defensive Team. For his career, he averages less than 1 steal and 1 block per game, and has never posted above league average PER. Defensive skill is often graded on a more subjective scale than offense so it can be difficult to quantity defense. Nonetheless, we use PER in this paper because it is one of the most well known advanced statistics, and will conflate PER and player quality henceforth.

2.3 Previous Works

Several past studies regarding the contract year effect in the NBA have been done. Gaffaney uses traditional OLS with controls for various subgroups of players and box score statistics and suggests that there is actually evidence of under-performance in certain per 36 minute statistics [1], although his usage of only 2012-13 season data large concern as it represents a fairly small sample size. White and Sheldon use motivation theory to estimate the effects over three year periods starting the year before a contract year and ending the year after [2], getting mixed results. Berri and Krautman focus solely on the effect of leaving a contract year and entering a new contract, and also find mixed results [3]. Ryan uses a two-stage least squares instrumental variables framework and finds a positive effect in win shares per 48 minutes for players in their contract years and claims that while there is an effect, national media typically overstates the contract year effect [4]. Ryan also focuses more on carefully defining contract year than previous studies, something we use as a guideline to build off on.

A majority of studies to this point boil down to looking at the average of some statistic for players in contract years vs players in non-contract years and taking the difference of those two as the estimated treatment effect of being in a contract year. We instead hypothesize that looking at year-to-year changes provides a better measure for estimating the contract year affect, which we detail in Section 4. In addition, we also develop a more stringent definition of contract years than previous studies to remove any potential biases in our estimate as we focus more on obtaining an accurate estimate of the contract year effect on a smaller subset of players than on producing a result generalizable across all players and contract situations. The exact criteria used to determine contract year players is defined in the next section.

3 Data

3.1 Data Collection

We collect player-year data for 6 NBA season ranging from the 2012-13 season to the recently completed 2017-18 season. Contract year status and yearly player statistics are not recorded together online. Traditional and advanced statistics for players on a yearly basis is readily available on basketball-reference.com. However for contract years, we had to manually code them into our dataset by using basketball-reference.com's NBA Player Contracts data (along with archives of the page from previous years), which is shown in Figure 1. There were well over 3,000 data to input by hand over the course of this data collection process.

We can see that Stephen Curry has an expiring deal in 2021-22, while Lebron James has

2017-18 NBA Player Contracts

2017-18 Salary Cap: $99,093,000 (Salary Cap History)

Average Salary: $5,752,807
Median Salary: $2,328,652

Largest Guarantee: Stephen Curry, GSW ($201,158,790)

Color Key: Free Agent, Player, Team, or Early Termination option

582 Contracts Share & more ▾ Glossary

			Salary							
Rk	Player	Tm	2017-18	2018-19	2019-20	2020-21	2021-22	2022-23	Signed Using	Guaranteed
1	Stephen Curry	GSW	$34,682,550	$37,457,154	$40,231,758	$43,006,362	$45,780,966		Bird Rights	$201,158,790
2	LeBron James	CLE	$33,285,709	$35,607,968					Early Bird	$33,285,709
3	Paul Millsap	DEN	$31,269,231	$29,730,769	$30,500,000				Cap space	$61,000,000

Figure 1: basketball-reference.com contract status data

a player-option for 2018-19. Therefore for Stephen Curry's 2017-18 season, would be considered a control year since his deal is not expiring. And even though it is widely known LeBron James plans to not pick up his player option for 2018-19, since 2017-18 precedes an option year, it is also considered a control year, but not a valid control year (more on this next).

3.2 Defining Valid Contract (and Non-Contract) Years

Recall that we above mentioned our dataset will measure year-over-year change in PER. This means each row in our dataset has two years worth of data. Therefore in order for a row to have received the contract year treatment, the players' previous year must have been considered a valid non-contract year and current year must be a valid contract year. Rows in which the players go from being in a valid non-contract year to another valid non-contract year are considered to be control rows. The following criteria must also be satisfied for both the previous and current year in order for that particular player year to be considered as part of our dataset (treatment or control).

- *Minimum of 10 minutes played per game and 41 total games.* Satisfying this criteria ensures that we a reasonable measure of a player's performance for both of those years.
- *Been among the top 300 in salary for any of the seasons in our dataset.* This prevents us from analyzing less talented players and dealing with rare exceptions to contract norms.
- *Provide both a treatment and control row.* Take for example players that only sign two year contracts. Each row that they provide will be a treatment row (rows for players going from a contract year to non-contract year are removed), since their previous year would be considered a non-contract year and current year would a contract-year. Since players that sign two year contracts are inherently different from players that sign three, four, or five year contracts, this would bias our contract year effect estimate

Metrics	**Mean**	**Median**	**SD**	**Min**	**1st Q.**	**3rd Q.**	**Max**
PER	15.23	14.90	4.31	4.90	12.30	17.60	31.50
Age	27.11	27.00	3.83	20.00	24.00	29.00	40.00
G	70.11	73.00	10.10	41.00	64.00	79.00	83.00
MP	26.24	27.00	6.75	10.00	20.50	32.00	38.70

Table 1: Summary Statistics

and these players should not be included in our dataset. Note that by design of our treatment and control rules, one year contract player-years are automatically excluded.

- *Both previous and current year are not considered option years.* This is the condition that reduces the size of our sample the most, due to the prevalence of options in the NBA. Take for example LeBron's 2017-18 contract status in Figure 1. Since it precedes an option year, we do not know with complete certainty whether or not LeBron is treating this current year as his contract year, or if he plans to pick up the option and stay with the team another year. Even if it turns out that he declines his player option, it should still not be considered a valid contract year because he was aware he had the opportunity of a 'second final year' of his contract if the first one did not play out well. True contract year players should not have any uncertainty in their contract situation upon the completion of the current season. They should know if it is their last year of their contract before the season begins. Team options for rookies are a special case in that they are typically picked up at the beginning of the prior season, so these prior seasons would be valid non-contract years (not an option year), since the player knows for sure they are guaranteed a contract the following year despite it being a year that technically precedes an option year.

3.3 Exploratory Data Analysis

The original NBA Player Contracts Data from 2012-2018 included 1060 players. After taking into account only players that were ranked in the top 300 in salary for at least 1 season, we have 2114 rows with 450 unique players. Applying the remaining criteria mentioned in Section 3.2 leaves the reduced sample with 865 observations and 336 unique players. It should be noted that this reduced sample will be further truncated in Section 4 for model specification, thus only a select number of players will be included in the modeling stage.

Table 1 reports summary statistics of the outcome variable of interest PER and some of the relevant factors in the data cleaning criteria, including age, games played per season (G), and Minutes Played per game (MP). PER is a normalized annual index that is mechanically centered around 15 and typically ranges from 5-30, so we can immediately see that our reduced sample still contains a wide range of players of varying skill. Interestingly enough, even though we have selected the top 300 players with regard to salary, which constitutes the top 32% of the players played during the 6 seasons, the median and mean, instead of reflecting an average PER higher than the league average (15.00), is only respectively 14.90 and 15.23. This suggests there is a fair amount of under-performing players with respect to their salary.

4 Methodology

4.1 Identification Strategy

Since we are working with observational data, we use the selection on observables assumption to identify ATT and ATE. Selection on observables comes in two parts. The first is conditional ignorability which states,

$$\{Y_{0i}, Y_{1i}\} \perp\!\!\!\perp D_i | X_i = X.$$

In the context of our data, this means that upon controlling for variables such as age, position, and tenure, we can consider whether or not a player is in their contract year to be as if random. In addition to conditional ignorability, selection on observables includes the common support assumption,

$$0 < \mathbb{P}(D_i = 1 | X_i = X) < 1$$

which is the assumption that we have a player in their contract year for each value of X. As we detail later, we perform matching with exact matching only on position, so this assumption is not a problem.

4.2 Methods Used

We ran two main types of models - our primary model uses matching as an estimation strategy and our secondary model uses regression. The primary model also includes three other related models that differ only in the construction of the outcome of interest.

In reading through the previous contract year analyses, we came to the conclusion that the outcome measure could be improved upon. Conceptually, measuring the impact of a contract year on PER seemed less pertinent a goal than measuring its impact on year over year change in PER. We are working with panel data, and the former largely ignores the relationship between consecutive player years that the latter hones in on. In addition, using the former bakes in an assumption that the effect of a contract year is additive, and we thought it more likely that larger shifts in PER might be seen in players with higher baseline PER levels. For these reasons, the outcome of interest is set as percentage growth. That is, for a player at time t, this growth outcome measure is $(\mathrm{PER}_t - \mathrm{PER}_{t-1})/\mathrm{PER}_{t-1}$. As mentioned above, three related measures of year-over-year PER change are included as well: growth in log PER, difference in PER ($\mathrm{PER}_t - \mathrm{PER}_{t-1}$), and difference in log PER. Our a priori preference for growth is due to its interpretability and non-additivity.

Our four primary analyses use matching to estimate the average treatment effect for the treated (ATT). To reiterate, we define treatment here as follows: using our definition of contract year above, a player in a contract year following a non-contract year is considered a treatment row, while a player in a non-contract year following another non-contract year is considered a control row. Any player not in either of these situations is not included as treatment or control. We choose to match on position (exact match), age, tenure (years in the league), as well as the t-1 year's PER, minutes played, and games started. We debated about the legitimacy of matching on year PER_{t-1} given that it is used to calculate our outcome of interest, but because it is very clearly a measure of something that happens temporally prior to the contract year, we feel comfortable using it here as a means to match similar players. The same is true for minutes played in year t-1, which is one of the components of PER_{t-1} and is thus present in the outcome measure, but is also useful to match on. The matching used bias adjustment as another means of controlling for differences in treatment and control units, and Mahalanobis distance given the wide-ranging scales and correlation of the covariates.

Given that, as described above, we are making three large changes from previous analyses in specifying our primary model - using a restricted subset of players, redefining the outcome of interest, and using matching rather than regression - we recognize that it would be unwise to compare our results directly to those previous results. However, in order to bridge the gap between the two, we attempt to mirror the standard analyses by running a player fixed effects regression model on our methodically constructed cohort. For this secondary model, the outcome of interest, as it has been for previous studies, is the yearly PER for an individual, and the quantity of interest is the average treatment effect of being in a contract year. We use age, age^2, and position, in addition to the player fixed effects, as covariates, closely mirroring Ryan's basic specification [4]. We expect the effects on the age covariates to suggest a concave relationship with PER.

In addition to running these models, we also run a placebo test as well as tests of the uniformity of our estimate across subsets of the data. We were not convinced that any data point at time t was unrelated to contract year, so we decided to use a lagged version of our outcome as a placebo instead. Due to some constraints stemming from the construction of our data (not all player years have lagged year-over-year outcome data, but every player year has lagged single-year outcome data) we could only run the placebo test with the regression model.

Given our focus on obtaining a homogeneous cohort in the interest of removing bias, we are interested in whether the effect estimate varies widely within subsets of our final cohort. We test this by calculating the ATT in our matched analysis for various quantiles of player quality (measured by PER_{t-1}) as well as for each of the positions, in effect measuring various forms of Conditional Average Treatment effect for the Treated (CATT). We used four quantiles of PER: three containing 30% of the data each, and a fourth containing the 10% of player years with the highest PER. We constructed them this way to isolate the effect for a small group of high-performing players we thought might have a different effect. Finding that little variation exists across these subsets would allow us to confidently claim our effect estimate is the true effect of a contract year without caveats for very high or low quality players, for example. Given that our four outcome measures of PER change listed above will have different amounts of such variation, we check, with the help of visual aids, if any of them are large or small enough to convince us for or against the use of a given outcome measure. One way of thinking about this is that we are imagining the Platonic Idea of a contract year effect that is constant across subsets of players, and checking if our outcome measures do a good job of recovering the form of that effect.

Table 2: Balance Table Before Matching

Treatment2	**PER**	**Age**	**Tenure**	**MP**	**GS**
Mean.Tr	14.288	26.935	5.755	25.166	39.516
Mean.Co	15.279	26.275	5.188	27.243	47.146
t-test p-val	0.014	0.114	0.165	0.001	0.011

Table 3: Balance Table After Matching

Treatment2	**PER**	**Age**	**Tenure**	**MP**	**GS**
Mean.Tr	14.288	26.935	5.755	25.166	39.516
Mean.Co	14.461	26.348	5.297	26.020	42.465
t-test p-val	0.689	0.196	0.306	0.218	0.373

5 Results

For our primary model, the results of our matching procedure is shown in balance tables 2 and 3. They suggest a success at balancing our matched covariates, though it is important to note that does not prove the treatment and control units do not differ on other covariates.

The treatment had positive effects on all four variations of our PER change outcome at the $p < 0.05$ level (Table 4). The growth outcome had the smallest standard error proportionally, though not significantly so.

Our secondary model found a statistically significant positive relationship between contract year players and PER (Table 5). The age variable effects confirm our expectation of concavity, and some significant differences between positions are observed as well.

Our placebo player fixed effects regression test initially showed a significant negative effect of being in a contract year at time t on PER_{t-1}, which was concerning. However, it quickly

Table 4: ATTs for four PER change outcomes

PER Outcome	**ATT**	**P-value**
Difference	0.691	0.038
Difference of logs	0.055	0.023
Growth	0.060	0.016
Growth of logs	0.023	0.020

Table 5: Player Fixed Effects Model

Variable	**Coef.**	**P-value**
Contract Year	0.515	0.010
Age	3.489	<0.001
Age^2	-0.067	<0.001

became clear that this effect was being driven by the subset of non-contract year players whose t-1 year was a contract year, boosting their PER_{t-1} values. That is, given that a positive treatment effect was observed, the construction of our placebo test was capturing that effect as part of the control group, resulting in a negative treatment effect. Once we removed that subset (effectively leaving us with exactly the same control cohort that was used as potential matches in our primary model), no significant placebo effect remained. It is worth noting that though this placebo test was conducted on a different group of player years than the true secondary model, we found a very similar relationship between contract year players and PER (coefficient= 0.487, p-value 0.029) when we reran that secondary model on the player years included in the adjusted placebo test. This leads us to feel confident in claiming a successfully effect-less placebo test.

Finally, our tests of uniformity gave mixed results. We calculated the CATT for our four PER quantiles and five positions, for each of the PER change outcome metrics. Given the different scales for our outcome metrics, we divided these effects by the overall effects for each metric to get a standardized spread of effects. This allows us assess whether any outcome metric shows much more or less variation than the others. The effects are plotted in figures 2 and 3, where each number represents one subset, and in which effects that did not vary between subsets would be clustered around a subset-to-overall ratio of 1 (the horizontal lines). In figure 2, the numbers in the plots represent the four PER quantiles in ascending quality order, and suggest that the better a player, the less we might expect a contract year to result in a bump in performance. (In fact, the top 10% of players are expected to see a *negative* effect on year-over-year PER change when going into a contract year as opposed to a non-contract year). We see much more effect size variation when measuring raw year-over-year PER difference as opposed to the other three outcome metrics. In figure 3, 1 and 2 refer to Point Guard and Shooting Guard, for whom our model finds little to no effect, while 3 and 4 refer to Small Forward and Power Forward, for whom our model finds about double the overall effect.

6 Discussion and Next Steps

Cumulatively, we believe that these results speak to our causal quantity of interest. We estimate that the causal effect of moving into a contract year as opposed to a non-contract year is a marginal increase in year-over-year PER of 6%. Our modeling strategies and various checks bolster our claims in the following ways:

- Our treatment and controls are homogeneous to a previously unobserved degree.
- Our matched groups are well-balanced on covariates of interest.
- The placebo test helps our claim that this effect indeed exists.
- Neither test of uniformity gave us reason to avoid PER growth in particular as our measure of choice.

In addition, this causal claim does not rely on the linearity assumption that is built into the commonly used player fixed effects models.

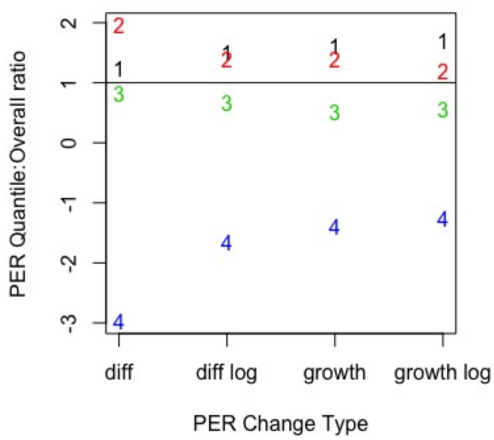

Figure 2: CATT by PER quantile and outcome measure

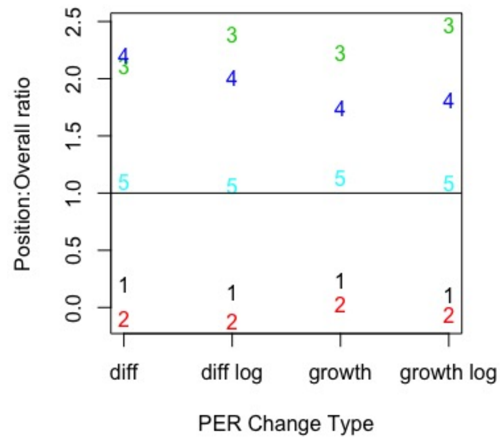

Figure 3: CATT by position and outcome measure

Our study raises certain concerns and had large limitations. Our primary concern relates to the widely ranging CATT estimates for different qualities of player, as shown in figure 2. Given the effort and focus we put into narrowing in on a specific type of player and situation being studied, it was disheartening not to see a near-constant effect size across subsets of our cohort. We are also concerned that there may be fundamental differences between players who get signed to contracts of different lengths. Estimating the effect size among players on a 2- rather than a 5-year contract would solve that issue (in addition to our completed breakdowns by player quality and position), but the sample size would drop greatly in such sub-analyses, and the data would be difficult to collect.

One limitation is that PER isn't a perfect measure of player quality, though as we explained in section 2.2, we consider it the best single-metric option. Another limitation is that we only had access to contract data over the past six seasons, and more data will only be available as future seasons play out.

A final potential issue is that we could imagine that there exists some unknown effect related to contract years that is experienced outside of the contract year itself. For example, entering the year before a contract year might also positively affect performance, which would mean our estimated effect size here is an underestimate. Alternatively, the year after a contract year may be treated as a year for relaxation, which, though boosting non-treatment PER growth, might also mean we are underestimating the true contract year effect. Other versions of the temporal spilling of the contract year's effects could potentially be affecting our estimate to an unknown degree and in an unknown direction.

One next step for this project would be to include controlling for or matching on overall team quality, perhaps using the team's win/loss record. Another would be to run through the same modeling process with other all-in-one player performance statistics. Finally, we could attempt to apply these methods to Major League Baseball contracts to investigate the effect of contract years across different sports.

We hope this paper, in addition to adding another estimate of the effects of the NBA contract year to the literature, serves as a thorough study in avoiding sources of bias in this space, as well as suggesting a new change-in-performance-focused outlook on the manifestation of the contract year effect.

References

[1] T. Gaffaney, "An analysis of the contract year phenomenon in the nba: Do players perform better or worse," 2016, Honors Thesis, UC Santa Barbara.

[2] M. H. White and K. M. Sheldon, "The contract year syndrome in the nba and mlb: A classic undermining pattern," *Motivation and Emotion*, vol. 38, no. 2, pp. 196–205, 2014.

[3] D. J. Berri and A. C. Krautmann, "Shirking on the court: Testing for the incentive effects of guaranteed pay," *Economic Inquiry*, vol. 44, no. 3, pp. 536–546, 2006.

[4] J. Ryan, "Show me the money: Examining the validity of the contract year phenomenon in the nba," 2015, Bachelor's Thesis, Harvard College.