Investigating salary change for NBA three-point specialists over time

Austin Hsieh

November 8, 2018

Introduction

Scoring in NBA basketball seems simple; either you take a shot outside the three-point arc for three points, or you take a shot inside the arc for two points. However, the NBA has evolved significantly as the years have gone by. Since the introduction of the NBA three-point line in 1979, three-point shooting has grown in importance as a skill for NBA players to develop (Fichman & O'Brien, 2018). And as the use of the three-point shot has proliferated, so has the use of three-point specialists: players whose primarily skill is shooting the three-pointer (Martin, 2013). But it's not immediately clear if our favorite three-point specialists are correspondingly getting compensated more, or if they're still getting paid the same as their two-point inclined teammates. Does the increasing importance of 3-point shooting correspond with an increase in pay for three-point specialists?

Literature Review (expand this)

First, I will look at why three-point shooting has become so popular. Is three-point shooting just a trend, or have teams and players found a more effective way to score more points? Mark Finchman and John O'Brien say the latter. In 2018, they analyzed three-point shooting in the NBA by treating it as an investment portfolio. They argue that a three-point shot is like a high-value, high-risk investment, and a two-point is like a low-value, low-risk investment. As a result of their economic analysis, they found that NBA players/teams take three-pointers at "nearly optimal" rate, which categorizes three-point shooting as an efficient game strategy, not just a popular trend among players (Finchman & O'Brien, 2018). It makes sense to then think about whether three-point shooters are being paid more, because teams want to compensate players that contribute to winning.

A 2013 article by Josh Martin looked into the emergence of three-point specialists in the NBA. Martin notes the introduction of entire offensive and defensive systems centered around taking the three-point shot. He

also found that in 2013, there were almost 20 players (out of a total of 500 in the NBA) he categorized as three-point specialists - players taking 60% or more of their shots from three-point distance. With a similar categorization, I can compare three-point specialists from many years.

Existing work has been done focusing on what skills generally gets NBA players paid more. A 2015 study of the 2013-2014 NBA season found that the most influential variables for pay were points per game, FG% (shots made divided by total shots), rebounds, fouls, and assists, in that order (Lyons Jr., Jackson Jr., Livingston), through multiple regression analysis. Unsurprisingly, this shows that players are paid primarily for how well they can score points. Three-point shooting falls under this category, as shooting threes affects both points per game and FG%.

My topic is a specific investigation of how player performance affects pay, but pay can also affect performance. A 2013 paper was conducted around player performance in contract years, or years where players are about to become eligible for a new contract. It found that players actually tend to perform more poorly across metrics like scoring and efficiency in contract years (Gaffaney, 2013). It may be more difficult to find a relationship between pay for three-point specialists and time if abnormalities like this drop in performance exist.

While these previous works provide thorough analysis of NBA data, none of them cover specifically whether three-point specialists have been paid more as three-point shooting has become more valuable. Lyons Jr. et al. determined that scoring was the primary factor (by far) for determining player salaries, but scoring is a broad and heterogeneous category, of which three point shooting is just a part. My scope is much narrower, focusing only on how three-point specialists are paid over time.

Method

There are two relevant datasets to this data analysis; the salaries of players and the stats of players. I pulled the salaries from data.world, and the stats from a dataset hosted on Kaggle. Both datasets were scraped from basketball-reference.com, a site that filters and retrieves a wide range of NBA statistics over a wide range of years. The salary data was available up to 2018, and the stats were available up to 2017. The integrity of this data is trustable, as basketball-reference is a widely used site for basketball statistics, and data.world and Kaggle are both well-known data science sites. Note that from here on, years will be referred as the year that ends a season. For example, when I say 2000, I mean the 1999-2000 season.

For the data analysis itself, I combined the NBA salary and stats datasets and filter that down to a subset of only three-point specialists. Initially, I wanted to filter the dataset to all years when the three-point line was present, but I chose to only include information from 1998 to 2017, because the NBA three point line was

shortened from 1994-1997.

Dealing with NBA salary is tricky, because of factors like salary "cap" changes (how much each team is generally allowed to spend on players) and inflation. I chose to normalize player salary against the NBA salary cap for that year, as this accounts for the generally increasing amount that teams allowed to pay players as well as inflation (because the salary cap is given in dollars from that year). I will refer to this new number as "Salary Ratio." As an example, if a player has a salary ratio of 0.01, then that player's salary would be 1% of the maximum allowed pay for all the players in the NBA. The NBA has what is called a "soft" salary cap, which means it can be exceeded, but this only exists in certain cases, most of which do not apply to three-point specialist type players anyways. Additionally, while these exceptions allow for salary raises, these raises are small in comparison with the total salary cap (Coon 2018).

I adjusted the criteria for "three-point specialist" to make sure the set of players was usable (for example, not too many or too little players), performing exploratory data analysis on both the subset and superset to look at general patterns and potential problems or abnormalities. I ultimately used a filter of having 3PAr (proportion of shots taken that are three-pointers) greater than 50%, three point percentage above average for that year, and three-pointers made above 82. 3Par above 50% gave me the best sample to work with, as anything higher would result in almost no players being classified as three-point specialists in the 90s and early 2000s, but anything lower would result in too many players being classified as three-point specialists. I chose to filter players that were above average three-point shooters, because players taking a lot of three pointers but missing them are probably less likely to get paid (potentially confounding my analysis). Finally, making more than 82 three-pointers a game is a standard criteria for any data analysis involving the three-point shot ("Statistical Minimums").

With the data prepared, I can then attempt to find a relationship between year and salary within this subset of NBA players, using linear and polynomial regression. I will compare the models and determine which one is most effective (or if neither of them is) by examining the p-value of the year variable (how statistically significant is the effect of year on salary) as well as the coeffcients/beta values. Based on these p-values and coefficients, I will finally conclude whether pay for three-point specialists has changed over time.

Results

The first step of constructing a relationship between pay and time for three-point specialists was to perform exploratory data analysis. Given the criteria mentioned above, these were the number of three-point specialists for each year:
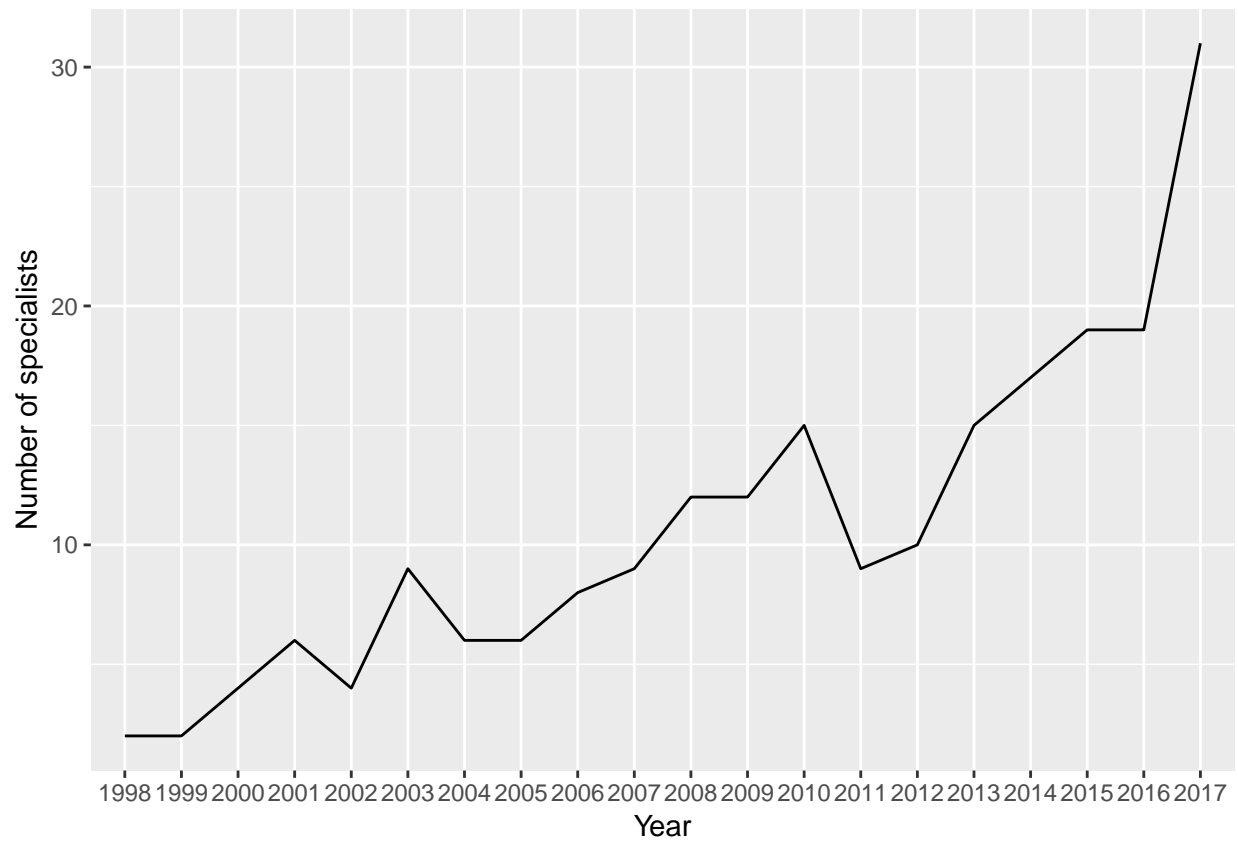
As expected, the number of three-point specialists has increased drastically over the years. This aligns with the increasing popularity of the three-point shot. However, even in earlier years, there were still three-point specialists (the minimum was two in 1998 and 1999). 2017 holds the peak in specialists at 31.

Another visualization I explored was the distribution of salary ratio for all three-point specialists. I created a histogram of salary ratio to look at this.
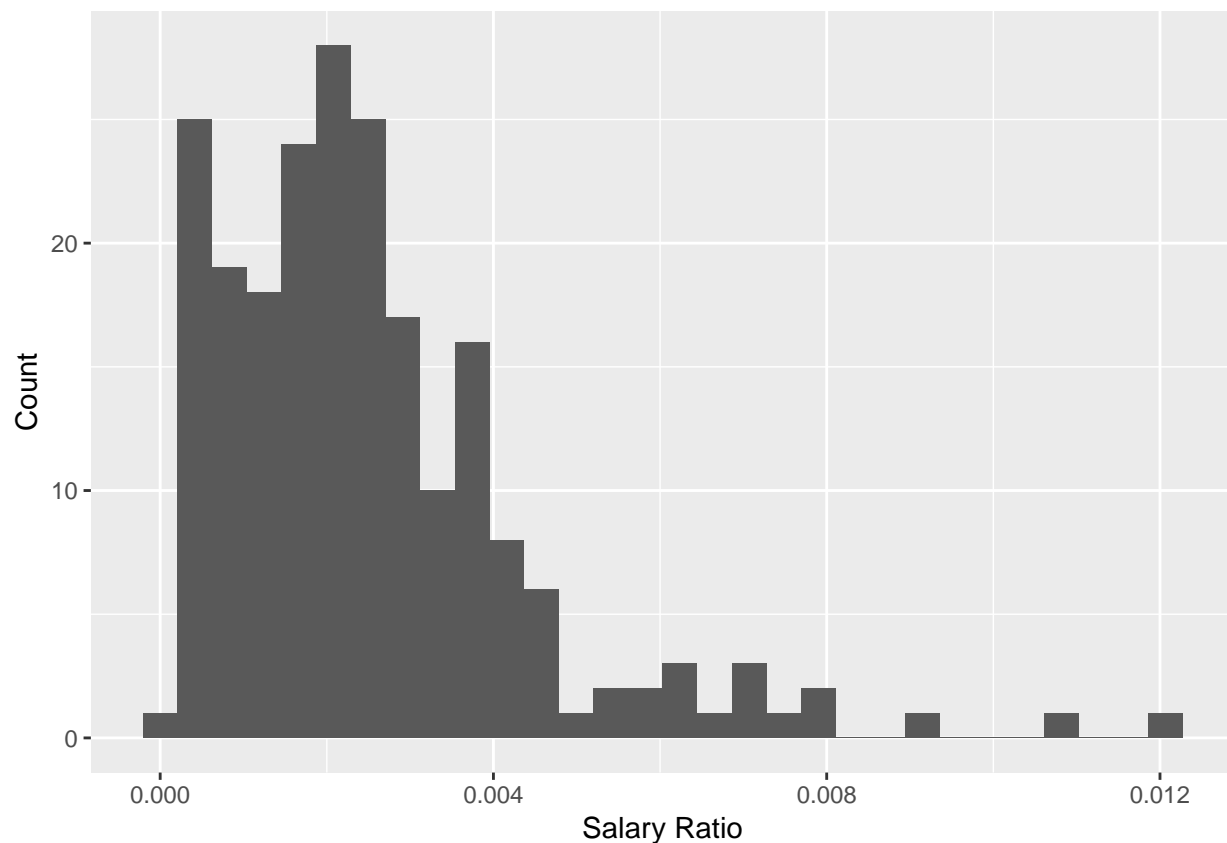
We can see that the distribution is skewed right. This is as we expect, because most players are average and get paid averagely, but a few star players will get paid more. Most below average players would probably not meet my criteria for being a three-point specialist, so they do not show up here.

I plotted the salary ratio (adjusted salary) of each player against the year that he played. This is the relationship that I'm attempting to model.
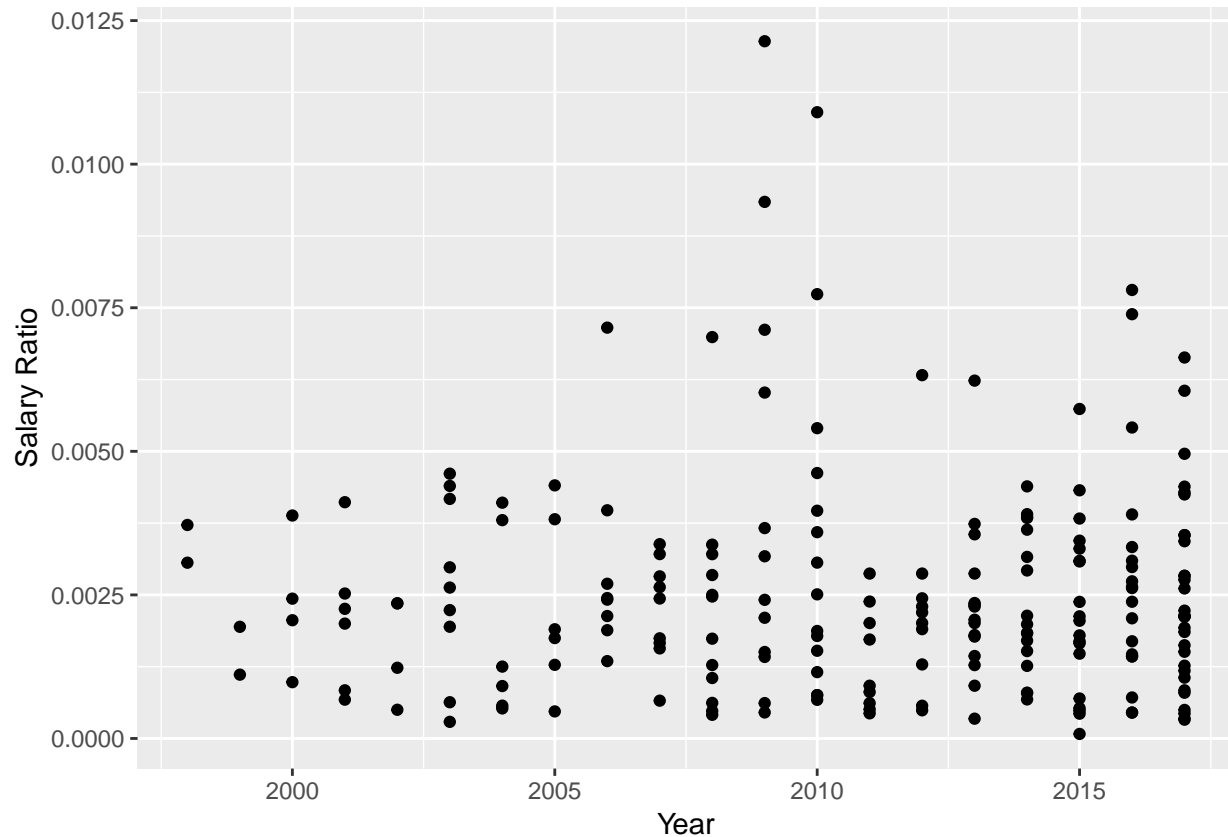
Looking at this scatterplot, it's not immediately obvious that there's any correlation between year and salary ratio. In fact, there's a spike of high salary values in 2010, which would reduce the strength of a linear regression. Nonetheless, I will try both linear and polynomial regression.

References

Coon, L. (2018, July 1). NBA Salary Cap. Retrieved from http://www.cbafaq.com/salarycap.htm

Davis, C. (2017, December 04). NBA Salaries - dataset by datadavis. Retrieved from https://data.world/datadavis/nba-salaries

Fichman, M., & O'Brien, J. (2018). Three point shooting and efficient mixed strategies: A portfolio management approach. Journal of Sports Analytics, 4(2), 107-120. doi:10.3233/jsa-160154

Lyons, R., Jr., Jackson, E. N., Jr., & Livingston, A. (2015). Determinants of NBA Player Salaries. The Sport Journal. doi:10.17682/sportjournal/2015.019

Gaffaney, T. (2013). An Analysis of the Contract Year Phenomenon in the NBA: Do Players Perform Better or Worse. Claremont McKenna College.

Goldstein, Omri (2018, April). NBA Player Stats. Retrieved from https://www.kaggle.com/drgilermo/nba-players-stats

Martin, J. (2017, October 02). Three-Point Specialists Rule the NBA. Retrieved from https://bleacherreport.com/articles/1559265-three-point-specialists-are-ruling-the-nba

Statistical Minimums. (n.d.). Retrieved from https://stats.nba.com/help/statminimums/