CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Clemson University: TigerPrints

Clemson University
TigerPrints

All Theses | Theses

12-2014

The Determinants of the Salary in NBA and the Overpayment in the Year of Signing a New Contract

Nuoya Li
Clemson University, nuoyal@g.clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/all_theses

Part of the Economics Commons

Recommended Citation
Li, Nuoya, "The Determinants of the Salary in NBA and the Overpayment in the Year of Signing a New Contract" (2014). *All Theses*. 2037.
https://tigerprints.clemson.edu/all_theses/2037

This Thesis is brought to you for free and open access by the Theses at TigerPrints. It has been accepted for inclusion in All Theses by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.

THE DETERMINANTS OF THE SALARY IN NBA AND THE OVERPAYMENT IN THE YEAR OF SIGNING A NEW CONTRACT

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts
Economics

---

by
Nuoya Li
December 2014

Accepted by:
Dr. Chungsang Lam, Committee Chair
Dr. F. Andrew Hanssen
Dr. Jaqueline Oliveira

ABSTRACT

Based on lots of previous research, the salary in National Basketball Association is determined by both personal characteristics and on-court performance. However, which kind of performance seems more important? Furthermore, does signing a new contract have any incentive or effect on the player's performance which makes the player be overpaid or underpaid in the year of signing the contract? And which kind of contracts and what kind of players tend to be overpaid?

In my paper I introduce the on-court performance, personal characteristics and salary. Then I build a connection between them and run 2 regressions to analyze the determinants of salary and overpayment in the year of signing a new contract.

In my first regression, I found that the player's salary is not just determined by the performance in the contract year, but the performance in the contract year and one year before that. I also found some interesting results like height is more important than weight, All star player salary premium exits in NBA. In the contract year, hitting a 3 point shot could always bring more expected future income than hitting a 2 point shot or hitting a free throw and some other funny results.

In my second regression, I found that offering a big contract always lead to overpayment in the year of signing the contract. I also found that providing a contract to an all star player could brings benefit to the team in the year of signing a new contract, and providing a contract to an old player might not be a wise choice.

The first regression tells us what factors are the determinants of expected salary, and which kind of performance is more important. It could help players to make the optimal choice of seeking a new contract. The second regression tells us what type of contract and what kind of players tend to be overpaid or underpaid. It may help team managers to make the choice when they provide a new contract.

ACKNOWLEDGMENTS

I would like to thank my advisor, Professor Chungsang Lam, he encouraged me to pursue a topic which I am really interested in and provided lots of insights and ideas to guide me finish the paper. I would also like to thank Professor Andrew Hanssen and Professor Jaqueline Oliveira, who helped me improve the theoretical model and empirical analysis. Finally I want to thank a Ph.D student, Andrew Swanson, he spent hours reading and discussing the paper with me, and gave me lots of useful ideas during the process.

TABLE OF CONTENTS

Page

TITLE PAGE i

ABSTRACT ii

ACKNOWLEDGMENTS iii

CHAPTER

I. Introduction and Literature Review 1

The Labor Market and Rookie Contract 1
The Feature of NBA Contract 3
Shirking Behavior and Contract Year Effect 7
Personal Characteristics which may influence the Predicted Income 9
The Relation between Performance and Salary 12

II. Data 16

What kinds of Variable I would use 16
Where I found the Database 17
Summary of the Data 18

III. Model 21

The Model I would use for the 1st Regression 22
The optimal Model for the 1st Regression 23

IV. Regression for the 1st Model and Analysis 27

Analysis for Personal Characteristics 28
Analysis for the Ability of Scoring Points 30
Offensive Rebound and Defensive Rebound 35
Analysis for Assist 37
Analysis for Block and Steal 40
Analysis for Offensive Performance and Defensive Performance 42

V. Regression for the 2nd Model and Analysis 46

The Introduction of Variables in the 2nd Regression 46
Summary of the Data 47
The Model I would use in the 2nd Regression 47

The Result of 2nd Regression 50
Analysis for Contract Length 50
Analysis for Sum of the Contract 51
Analysis for Age 51
Analysis for All Star 53
Analysis for On-Court Performance 53
Conclusion 54

VI. Weakness of the Paper 55

REFERENCES 57

CHAPTER ONE

Introduction and Literature Review

In this part I will build a connection between the “salary (contract)” and “personal characteristics and on-court performance”. In order to do that, I will find out what kind of contract and which year performance I should use. Because the NBA player salary (contract) is determined prior to the season, all of the contracts should be treated as a prediction of their future performance. The rookie contract is a prediction based on the performance in NCAA or other leagues. That is why we cannot use rookie contract to build the connection. Furthermore, Based on the features of NBA contract (multi-year, guaranteed with high level payment), the player’s optimal choice is to raise their effort in the contract year then shirk their behavior, which makes their future performance hard to predict. In the same time, the player’s personal characteristics like their nationalities and popularities might earn money for the team, which means these personal characteristics may also influence the player’s salary. In the end of this part I will introduce the main source of NBA income and revenue sharing policy, it would help explain why the NBA player salary is correlated with their performance and personal characteristics.

1. The labor market and rookie contract in NBA

Previous research shows there is a correlation between salary and performance in sport league. However, which kind of salary and contract we should use to build this connection? In this part, I will introduce the rookie contract and show the reason why we cannot use rookie contract to connect the on-court performance.

The NBA contract is determined prior to the season, which means the players’ current salaries have been determined before the current season. In other words, all of the contracts are the prediction of the future on-court performance.

According to “Are top draft picks overpaid?” by Nikhil Joshi (2011), the result shows that teams could extract significant benefits from the new players in NBA by their rookie contract restrictions. Obviously, the salaries of the rookie contract did not match the players’ performance.

The draft pick in NBA includes 2 rounds, each round includes 30 rookies (rookies means the new players without NBA experience). Since 1995, the salaries of the 1st round rookies have been determined by a certain pay scale based on each one’s draft order. Generally the 1st round rookies would get a guaranteed 3-years contract. In the same time, the higher the player’s draft order, the higher the salary he would get. For instance, the 1st pick rookie would get much more than the 20th pick rookie. In contrast, 2nd round rookies are not guaranteed with a certain scale of payment in their rookie contract, and they always get smaller rookie contract compared with the 1st round pick players. Besides the drafted rookies, there are also some undrafted rookies in every year, they would get an non-guaranteed contract with shorter length and smaller scale compared with drafted players.

Because rookie represents the player without NBA experience, the scale of rookie contract just depends on the draft order. And the draft order might be determined by the prediction based on his performance in NCAA or the foreign basketball leagues in Europe or Asia. It makes people hard to predict their future performance because the rookie contract is not a prediction based on NBA performance. That is why the rookie contract is always hard to match their on-court performance in NBA.

However, after the contract year of rookie contract (the year before the player could be an unrestricted free agent or a restricted free agent), the player could sign a new contract, and the new contract will generally match their performance well. That is the reason why some players’ salaries got dramatically increase or decrease after the first 3 or 4 years. Because after rookie contract, the game masters will have a more rational prediction about the players’ future

performance based on their previous stats in NBA. Overall, the contract after the rookie contract should be the contract we use to build the connection with performance, because that kind of contract is a prediction to the player future performance based on their previous performance in NBA.

2. The features of the contracts in NBA---multi-year, guaranteed with a high level payment

From the introduction of rookie contract, we know that when we try to build connection between contract and performance, we should use the contract after the rookie contract. However, based on the features of NBA contract, the multi-year guaranteed contract will have a strong financial incentive to the players' behavior and makes the players' performance hard to predict once they sign the contract. We call that phenomenon "contract year effect". The contract year effect will help us to find which year performance we should use to build connection between salary and performance.

I will introduce the features of NBA contract in part 2 and introduce the result caused by the features just like shirking behavior and contract year effect in part 3.

[1]Why the average salary level in NBA is higher than other leagues.

According to an investigation by Forbes website in 2012, the difference of average income among the 4 major American sport leagues was really significant. For National Basketball Association (NBA), the player's average salary was $5.15 million per year. For Major League Baseball (MLB), the player's average salary was $3.2 million per year. For National Hockey League (NHL), the number was $2.4 million per year. And for National Football League (NFL), the number was only $1.9 million per year.

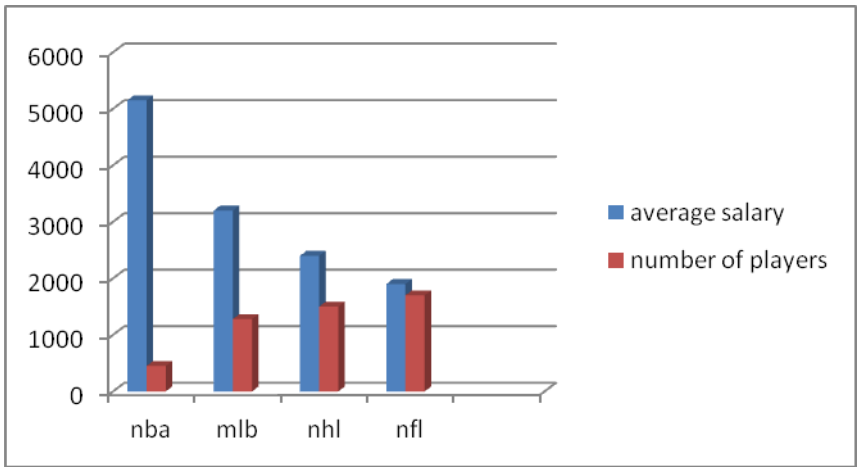

Figure 1.1 The average salary and number of players in the 4 major sport leagues in the US

Obviously, the salary level in NBA is much higher than most other major sport league based on the result of investigation, although the basketball is not the most popular sport, when we consider about the average attendance, the football(67604) and baseball(30514) are much more popular than basketball(17407) in the US.

In my opinion, the player's salaries are highly correlated with the team income, and based on the revenue sharing plan, most of the teams could earn some money, or compensate their loss. Then the level of the players' salary should be correlated with the number of the players in the league. NBA has the least number of players among the 4 major sport leagues in the US, that might be the most important reason why NBA players have the highest salary level among them.

For example, NFL has 32 teams, each team could have 53 players, which means it could be at most 1696 players in that league, that means the revenue (or sharing revenue) will be shared by almost 1700 players. Same situation are also exist in MLB, they have 32 teams and 40 players for each team, which makes the baseball league could have at most 1280 players. Or think about NHL, there are 30 teams, each team could have 50 players (23 are active players), makes the number of players in the National Hockey League about 690 to 1500.

However, NBA has only 30 teams and each team can only hold 15 players, makes the number of players in NBA 450 at most. The smallest group can share the most of the revenue, that is the main reason why NBA salary level is relatively high.

Based on the research of Neal Jean (2010), the author came up with an interesting idea. As what he claimed, unlike football, the vast majority of players play either offence or defense but not both, in basketball area a player has to contribute to both, we never heard which NBA player just focus on scoring the points but never defends, the player might not be a good defender (like Nash and Jeremy Lin), but it does not mean he could give up defending even for 1 minute.

Unlike base ball, most players only spend a fraction of the playing-court. The basketball player has to contribute to the entire court. The basketball players need to keep running and positioning all night on the court, never stop when they are on the playing court.

Each basketball team can only put 5 players on the court, which means the game will heavily rely on 1 or 2 core players' performance for each team. Unlike football or soccer, buying 1 or 2 super stars might not change the team. In NBA, getting a star player could totally change a team, it could let a terrible team which missed the playoffs last season become a potential champion competitor this season.

Overall, compared with baseball, football and soccer, the basketball player need to contribute more both on offensive and defensive parts for the team, and the wining relies heavily on a few players(5 to 8 major players), so that no surprising that NBA players earn more in average than other sport league.

[2] why do not the teams use partial-guaranteed short contract to replace guaranteed multi-year contract

Based on the research of Neal Jean(2010), the contract structure in NBA is significantly different from MLB and NFL. For instance, most of the contracts in NFL, the large proportion of

the contracts are not guaranteed, the players can be fired or released at the team's discretion. However, most proportion of each contracts are guaranteed in NBA, especially for the long contract.

Since most of the contracts in NBA are multi-year and guaranteed, players optimal choice may be to raise their performance in contract year and reduce the effort during the next contract. Without a doubt, the boss and managers must notice that, they have already known what would happen once they offer a contract to the players. Actually the team managers have incentive and reason to do that. The following reason could help us understand that.

① Guaranteed contract is good for the players

Stiroh(2007) has discussed about why long term contract are not replaced by the incentive-based (partial guaranteed) contract or short term contract. Stiroh gave some reasons to explain the benefit of guaranteed contract for stimulating the players. First, Playing time is decided by coach and the team strategies, the players cannot even get enough time to perform if the coach let him sit on bench if the teams use partial-guaranteed contract or incentive-term contract. Second, players tend to avoid of uncertainty, like career-ending injury. Guaranteed contract can motivate them to play harder compared with partial guaranteed contract. Based on Neal Jean's opinion, if we use incentive based contract, we might have to dramatically increase the average salary level, it might be more expensive than current situation.

②Guaranteed contract is good for the teams

Based on the study of Prendergast (1999), multi-year contract is much better than short contract. Different with Stiroh, Prendergast stated his opinion in the team's stand point. He provided several reasons to show that multi-year guaranteed contract has more advantages for the team. First, players learn to play more effectively with specific teammates. Then they can learn how to play smart with certain teammates and win more games. Second, create fans and develop

loyalty with supporters, then increase box office revenues. Third, players also tend to choose longer contract in terms of avoiding the career ending injury and high level uncertainty. If the teams use short contract to replace long contract, they might pay higher per year salary.

In conclusion, based on previous studies, compared with a short non-guaranteed contract, providing a multi-year guaranteed contract is better for both players and teams.

3. Shirking behavior & contract year effect

As we analyzed above, most contracts in NBA are guaranteed multi-year contracts, the managers and bosses do that is because that kind of contract could bring benefits to both teams and players. However, it may also bring "shirking behavior" and "contract year effect".

Based on the previous research, shirking behavior is always correlated with contract year effect. Shirking behavior(Alchian and Demsetz 1972 Holmstrom 1979) includes: Increasing in effort expended by the agent before signing a new contract and decreasing in effort expended by the agent after signing a new contract. Strategic behavior as a result of long-term guaranteed contract always happened in the real world. When the agents face with the incentives created by long-term guaranteed contract, they will have motivations to vary their efforts to maximize their own utility.

Based on a research of Kenneth Lehn in 1982 about the disabilities in the baseball league, he found a lot of players spent more time on the disability list since 1977, which means they spent less time on the playing court. During that time, a lot of people believe it is because the new material of the turf has increased the risk of getting injured, or because of the expansion of baseball league in 1977 made a lot of younger players join in, and younger players had no experience so they were easy to get injured. However, based on the regression analysis of Lehn, he found that the new turf and expansion of the league did not influence the players' behavior too

much. The most significant reason is because the multi-year guaranteed contract has been widely used since 1976, which might lead the players change their behavior, because most of their future income has been guaranteed, the player might change their behavior after signing a new contract.

That paper of Kenneth Lehn is one of the earliest papers which analyzed the relationship between the guaranteed contract and the player's behavior in the sport league, it could be treated as an evidence that the shirking behavior always exist in the sport league since the guaranteed contract appeared. After signing a multi-year contract with most proportion guaranteed, the players have reason to be less hard working than before. And in order to sign a multi-year guaranteed contract, the players have incentive to player much harder in their contract year than any other time.

The "contract year" means the year before the player becomes a free agent (restricted or unrestricted). A free agent could sign a contract with any team who provides the most attractive contract. In order to earn a higher-paid contract for next several years, the players will have incentive to play harder or more selfish to increase their on-court performance. After getting a new multi-year guaranteed contract, the player's performance always decrease a lot compared with their performance in the contract year. Then we call this phenomenon as "contract year effect".

According to the research of Berri and Krautmann, there exists the "effect of contract year" which means the players tend to perform harder in his last year of the contract before signing a new contract. The change of the players' behavior before and after signing a new multi-year contract implies that, the only solid connection between the salary and performance is "the average salary of the new contract" and "the performance in the contract year and the performance before that".

4. **Personal characteristics might also influence the expected income**

-----Nationality, All Star and Age. Three factors out of the playing court which may influence the expected salary

A lot of reports from sport website talk about some players are overpaid and some are underpaid, some of them got much more than what they should have based on their current and previous performance when they signed a new contract. Actually, the players might not be overpaid or underpaid. The salary premium may be caused by some personal characteristics. From the analysis above we know that on-court performance may influence the player's expected salary, now I want to talk about some personal characteristics which may influence the players' expected salary, and I believe that characteristics should be correlated with the ability of "making money for the team". What kind of personal characteristics could represent the ability that "making money for the team out of the basketball court"? I suppose there are at least 2 main factors. Nationality and All Star.

[1] Nationality

In the article "Does the salary discrimination exist?" by Min-Hua Kuo and Wen Jhan Jane and Sheng Tung Chen, the result indicates that a significant salary premium of international players exist in the Base Ball League. I am interested about that, I also test the effect of "being an international player in NBA". No surprising, my result showed that the premium payment for international players does not just exist in baseball area, but also exists in basket ball area.

Think about Yao Ming, the "Great Wall" from China, which entered in NBA in 2002 as 1st pick of 1st round. Yao became an excellent center later in NBA, however, he had struggling time, (he only got 10 points in his first 4 games, 2.5 points in average, this shameful record was broken by Bennett in 2014). Surprisingly, the Rocket gave him a huge contract before his rookie contract finished. And then Huston Rocket tried countless times to sign an extension new contract with

Yao even the current big contract is not over. And I have to say, the contract is too large(6 years 130 million in 2009), much higher than the other centre who had similar performance in the same time. Although this contract did not come true because of Yao's career ending injury on his knee, it is still valuable to analysis. Why could Yao get that fat contract? At least one reason is obvious. He represented the market of a country that has 1.3 billion people, a huge potential market.

In my opinion, the reason that international players tend to get premium salary even their performance are not as good as domestic players is not because of their nationality, but the potential market behind the nationality. It reminds me that when Yi, another Chinese guy entered NBA, he threatened the Milwaukee Bucks to trade him (this action was impolite I have to say), his reason was "I represent 1.3 billion people", I suppose he knew exactly his commercial value, which gave him the confidence to stand on the opposite side of his team.

[2] All Star

The other factor is All Star. This is really crucial. In the research "Does the All Star Premium exist in the NBA(2006)" by Jame Hayles, he found a strong evidence to prove that all star players tend to have a premium salary than normal players. And in my research, I also add the "all star" into my regression. In his research he built All Star like a dummy variable, different from what he did, I focus on the influence of how many all star the player has played before he signed the new contract. So that in this way I can see more details of the effect of being an All Star, after all, the effect of being all star for once and being an all star for ten times is totally different before signing a new contract.

A lot of people believe the all star players are good player. It might be, it might not be. As I know, a lot of good players never became an All Star player during his career so far, like Monta Ellis, the guy who could get 25 points and 5.3 assists per game in 2009-10, with 19.4 points and 5.0 assists in his career. Or Josh Smith, the guy who can get 15.5 points, 7.9 rebound in his career

so far, with the best season 18.8 points 9.9 rebound, never become an all star neither. Every year there are more than 400 players playing in the league, anybody can make sure the 24 all star players are the top 24 players in NBA? Nobody could, because All Star players are not “good” players but “hot” players.

First we need to be clear how the All Star players be chosen. The 5 started All Star players for each team are chosen by fans, they vote online to choose the players they want to watch. (The NBA started to offer the Spanish and Chinese language for fans to vote since 2003). The 7 bench All Star players are chosen by a committee of coaches. That is why the All Star players do not represent “the best performance” but “the highest popularity”.

Consider about Yi, again, he got the top 3 votes in 2008 in the forward position among all the players of 15 eastern teams, but his terrible performance was not qualified to be an All Star. This example tells us All Star players do not always represent the best players. All Star means the guy is popular, people wants to watch the guy’s personal show, when the audience want to watch someone, they are willing to spend money on that, all star players represent the ability of earning money somehow. When they get the ability to attract the audience, they get the power of earning more money for the team, then they might get extra payment based on their popularity.

[3]Age

In sport area, the age and experience are really important. However, we all know that the experience is highly correlated with age. For most of previous research, they just take one of them. In this paper, I prefer to use age as the variable.

In my research, I just use “age” but not “age and age square”. Some researchers believed that the expected salary should increase in a decreasing rate with age. However, as my data showed, in the real world, salaries of old players (especially after 30) are always much lower than the younger players. Young players have potential to be better, to get progress and they have less

injure history. But the old players have no room to get progress, they have already become a certain type of player, they have had their certain role. Furthermore, old players' health condition and physical quality are decreasing year by year, especially in sport area.

Someone may notice that some old players can still get the contract with high level salary, like Garnett, Kobe, Wade and someone else. Then these samples give them an illusion that players' salaries should increase as they grow older. Actually it might be true, but I do not agree with that. Some old star players could get higher salaries is not because they are old, is because they could still get 20 more points or take 10 more rebounds per game. Or they are super popular outside the basketball court that they can attract audiences. Anyway, no one will pay anybody just because he is old, the teams pay the players high salaries is because they can perform well and earn more money for the team. Interesting, if we look at the data carefully, we always find a truth, the older the players, the shorter the contract length will be, even for the super stars. The team managers are always smarter and more rational than others.

5. The relation between performance and salary

An article which name is "Fans ought to walk out on overpaid greedy NBA" of Spokesman Review in 1999 came up with an interesting idea, that is the NBA players are overpaid because "The money they made in one season is more than a common man can see in the lifetime", even worse, because "American public keeps feeding their egos, then players believe they deserve the salary." I like the passion of this article, however, I cannot agree with the idea. Based on the assumption of rationality in economics, people will not hold a bad asset which will just lose money. I would rather believe the boss and team managers are more rational than common people because they will suffer a huge loss if they make a mistake about the contract to provide.

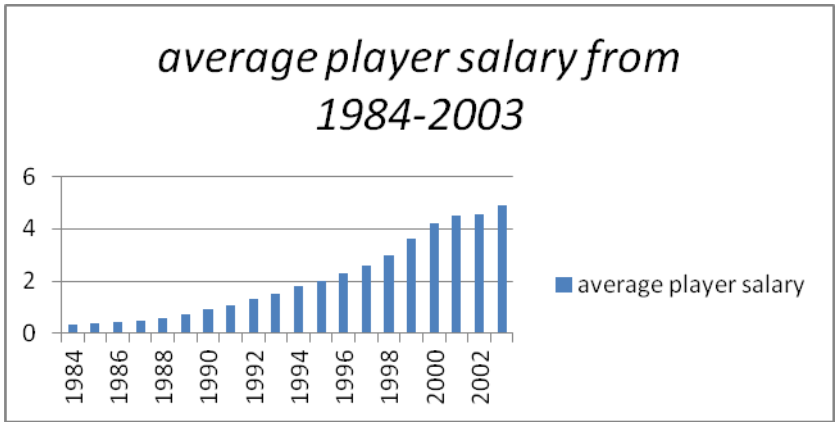

Figure 1.2 The Average Salary of NBA Players change in 1984-2003

According to a report in Hoopstuff, it mentioned some change in NBA for the past years. "Under David Stern's regulation, the amount of NBA team expanded by 25%, also added up 105 potential roster spots, which means more teams and more players." "The NBA expanded their brand internationally, increased the popularity of NBA to 215 countries". There are countless great accomplishments. Most important is "In 30 years under Stern, the NBA has increased its value by 3000%, mathematically means each team has increased its value by 100% annually." Think about Donald Sterlin, (the former owner of Los-Angeles Clippers) he sold the team for 2 billion bucks this summer, think about the amount of money, it might give us some hint about the profit the bosses would get in NBA.

We can conclude that the owners earned much more than before day by day with the development of NBA. That is why the players earned much more than before. Based on a report of "patricia's various basketball stuff", the average salary in NBA in 1984 was $0.33 million, it increased to $1.1 million in 1991, and $4.9 million in 2003. We know there is an inflation effect to the increase of salary, but we can still find out that the salary level in NBA has dramatically increased over the past 30 years.

We can find that the salary increase of the players and the profit increase of the team happened simultaneously. It tells us the relationship between player income and team income----

the player's salaries depend on how much they can make for the team, it is not because the American public feed their egos, it is not because the team managers are stupid, it is not because the players are greedy, none. It is just because the players can make more money for the team. This is how they get paid.

Then let us think about what kind of characteristics and what kind of performance could help team earn money. As what I stated in the beginning, nationality and All star experience might help team earn more because that kind of characteristics represent popularity and potential market.

Large proportion of the team income comes from the box office revenue, TV contract and licensed good sale like caps and clothes. Box office revenue and licensed goods revenue come from the fans who are willing to pay. What kind of factors could attract the audience to buy the tickets? Generally, the team which can always win the game will attract audience, because nobody will pay for watching a terrible team who will always lose the game at home. How to win the game? Getting more points and losing less, it depends on the players' on-court performance. Or the team which locates in a big city can earn more because big city always represents big market, and big market attracts super stars, that is a reason why Boston and Los-Angeles can attract the star players to join in, because the star players have more chance to get extra income like advertising endorsement in big cities. So the audience may also pay for watching the star players' personal show. It also depends on the players' on court performance.

Overall, in order to increase the team revenue, the team needs popular star players or winning more games. No matter which way they choose, it depends on the performance of the players on court. This is why a lot of previous research use on-court performance and personal characteristics as the determinants of salary in NBA.

In a word, the players' on-court performance can increase the team revenue by winning the game, and the players' personal characteristics (nationality and all star experience) might also increase the team revenue by expanding the potential market. These will make money for the team, and the team will pay for these abilities. That is why the salary is determined by their on-court performance and personal characteristics as a lot of previous research showed.

CHAPTER TWO

Data

I have collected 955 observations from 432 different players from 1995—2013, all of the players I chose had at least 2 contracts in the whole career, which means their salaries could be related with their performance.

What kind of dependent variable I would use

I will use the average salary of the new contract (except rookie contract) as the dependent variable, adjusted by CPI. (CPI comes from Federal Reserve Economic Data)

What kind of independent variables I would use

My independent variables include “personal characteristic” and “on-court performance” in the contract year and performance before.

On-court performance includes point, rebound, assist, block, steal, turnover and personal foul. To make it specify, point could be divided into 3 parts, 2 point, 3 point and free throw. And each one is consist with hit and miss. Similar, rebound could be divided into offensive and defensive. It makes sense, because the 2 kind of rebound have different functions.

Personal characteristics are some variables outside the playing court which can influence the players expected salary, like the players drafted or undrafted, age, height, weight, nationality and all star experience. Like we always believed, the tall guys have advantage in basketball area, now we can see if it is true by making a regression.

As what I stated above, I built a connection between the average salary of each new contract and the performances in different years before signing the new contract, including performance in

“contract year”, performance 1 year before contract year, performance 2 years before contract year.

Where I found the database

www.basketball- reference.com, this website provides NBA players salary and performance for each year

www.spotrac.com/nba/ this website provides some detail of the contract, like when the player sign a contract, how big the contract is and the length of the contract

basketball.realgm.com This website provides the contract and transaction information for each player, like when the player sign a multi-year or single year contract

USA Today, as a lot of previous studies showed, there was a database about NBA contracts in that website, however, I did not find the information and details there. The links about NBA contract seems like being canceled and all the other previous link did not work anymore, except the salary information for 2001-2002, this also provides a little bit help.

Introduction of the variables for the 1st regression

Dependent variable
Salary The average salary of the new contract, as annual salary, adjusted by CPI, use million dollars as unit
Independent variables
Draft Dummy, 1 if the player is undrafted, 0 if the player is drafted.
Height How tall the player is, use inch as unit
Weight The body weight of the player, use pound as unit
Nation Dummy, 1 if the player is international player, 0 if the player is domestic player
All star How many all star games the player has played until the contract year
Age The age of the player in his contract year
Hit2p How many 2 point shot the player hit (per game) in the contract year
Mis2p How many 2 point shot the player missed (per game) in the contract year
Hit 3p How many 3 point shot the player hit (per game) in the contract year
Mis3p How many 3 point shot the player missed (per game) in the contract year
Hit FT How many free throw the player hit (per game) in the contract year
ORB How many offensive rebounds (per game) the player got in the contract year
DRB How many defensive rebounds (per game) the player got in the contract year
AST How many assists (per game) the player made in the contract year
STL How many steals (per game) the player made in the contract year
BLK How many blocked shot (per game)the player made in the contract year
TOV How many turnovers (per game) the player made in the contract year
PF How many personal fouls the player made (per game) in the contract year
Hit2p_1 How many 2 point shot the player hit (per game) one year before contract year
Mis2p_1 How many 2 point shot the player missed (per game) one year before contract year

Hit 3p_1	How many 3 point shot the player hit (per game) one year before contract year
Mis3p_1	How many 3 point shot the player missed (per game) one year before contract year
Hit FT_1	How many free throw the player hit (per game) one year before contract year
ORB_1	How many offensive rebounds (per game) the player got one year before contract year
DRB_1	How many defensive rebounds (per game) the player got one year before contract year
AST_1	How many assists (per game) the player made one year before contract year
STL_1	How many steals (per game) the player made one year before contract year
BLK_1	How many blocked shot (per game) the player made one year before contract year
TOV_1	How many turnovers (per game) the player made one year before contract year
PF_1	How many personal fouls (per game) the player made one year before contract year
Hit2p_2	How many 2 point shot the player hit (per game) two years before contract year
Mis2p_2	How many 2 point shot the player missed (per game) two years before contract year
Hit 3p_2	How many 3 point shot the player hit (per game) two years before contract year
Mis3p_2	How many 3 point shot the player missed (per game) two years before contract year
Hit FT_2	How many free throw the player hit (per game) two years before contract year
ORB_2	How many offensive rebounds (per game) the player got two years before contract year
DRB_2	How many defensive rebounds (per game) the player got two years before contract year
AST_2	How many assists (per game) the player made two years before contract year
STL_2	How many steals (per game) the player made two years before contract year
BLK_2	How many blocked shot (per game) the player made two years before contract year
TOV_2	How many turnovers (per game) the player made two years before contract year
PF_2	How many personal fouls (per game) the player made two years before contract year

Summary of the data

Variable	Obs	Mean	Std. Dev.	Min	Max
salary	955	5.778373	5.364819	.4595071	34.96006
draft	955	.0397906	.1955693	0	1
height	955	78.98115	3.763677	66	90
weight	955	223.3812	28.48846	161	325
nation	955	.1591623	.3660191	0	1
allstar	955	.5549738	1.893901	0	16
age	929	26.55113	3.629059	19	38
hit2p	929	3.102906	2.055497	0	10.9
mis2p	929	3.342949	2.081687	.2	11.2
hit3p	929	.6651238	.6802557	0	2.8
mis3p	929	1.187191	1.117879	0	5.1
hitFT	929	1.916469	1.561111	0	9.2
misFT	929	.6287406	.5665635	0	5.7
ORB	929	1.142411	.8973434	0	5
DRB	929	3.093649	1.806493	0	10.7
AST	929	2.163402	1.922943	0	11.6
STL	929	.7562971	.4419364	0	2.9
BLK	929	.4966631	.5241985	0	3.7
TOV	929	1.408719	.7994606	0	4
PF	929	2.073305	.7068004	0	4.8
hit2p_1	917	3.047219	2.027401	.1	11.8
mis2p_1	917	3.333588	2.101685	.1	13.6
hit3p_1	917	.5983642	.640029	0	2.8
mis3p_1	917	1.105016	1.080513	0	5.1
hitFT_1	917	1.868484	1.51525	0	8.2
misFT_1	917	.6225736	.545108	0	5.4
ORB_1	917	1.132933	.8719209	0	5
DRB_1	917	2.996401	1.797347	0	10.8
AST_1	917	2.108506	1.937273	0	11.4
STL_1	917	.7394766	.4445337	0	2.8
BLK_1	917	.5074155	.5289538	0	3.7
TOV_1	917	1.383969	.7873691	0	4.2
PF_1	917	2.080698	.723927	0	4.4
hit2p_2	838	3.058473	1.967458	.1	11.8
mis2p_2	838	3.378401	2.022968	0	11.8
hit3p_2	838	.6107399	.6567573	0	3
mis3p_2	838	1.116229	1.094648	0	5.1
hitFT_2	838	1.890692	1.472637	0	8.2
misFT_2	838	.6321002	.5373355	0	6.4
ORB_2	838	1.135203	.869757	0	4.7
DRB_2	838	2.975895	1.766756	0	10.7
AST_2	838	2.081981	1.872685	0	11.6
BLK_2	838	.5089499	.5360403	0	3.6
STL_2	838	.7403341	.4261981	0	2.5
TOV_2	838	1.41778	.8026726	0	4.4
PF_2	838	2.099523	.7238717	0	4.1

.

Table 2.1 summary for the data in 1st regression

CHAPTER THREE

Model for the 1st Regression

Based on a lot of previous research by other people, “effect of contract year” exists. However, if everybody knows that, that means the boss and the team managers have already had a prediction about what would happen once they offer a big guaranteed multi-year contract to the free agent. Once we make a mistake about the model, maybe we need to rebuild a model and retest the hypothesis, it takes time. However, once the boss makes a mistake about offering a contract, it takes money, million dollars.

As a lot of previous studies proved, [1] the personal characteristics and [2] on-court performance are the determinants of the players’ expected salaries in NBA. However, as we analyzed before, the contract is not signed year by year and most of the contracts are multi-year and guaranteed, which makes the player could vary their efforts during the contract, it also means the players’ performance have nothing to do with their salaries during the contract.

Furthermore, based on the “contract year effect”, most players would increase their performance in the contract year, so the connection between “the performance in contract year” and “the average salary of the new contract” is most solid and reasonable.

According to the connection, our model should look like this:

salary = β0+βi*performance in contract year+βj*personal characteristics+ε

The model above is the model of some other research used.

However, does it mean the performance in other years before the contract years are not important? I do not think so. As the previous study, the NBA players tend to have a strategy behavior----which means raise effort before signing new contract and decrease effort once they get the contract. At the same time, I suppose the team managers also have noticed that, so when

they evaluate a play's value, the team managers might not just pay attention to the performance of the contract year.

Finally my 1st model is given as:

Salary=β0+βi* stats(i) +βi-1*stats (i-1) +……+ βi-t* stats (i-t) +β*chara+ε

The salary represents the average salary of the new contract.

Where stats(i) means the performance in contract year, stats(i-1) means the performance in one year before contract year, stats(i-t) mean the performance in t years before contract year.

This model will finally tell us before signing a new contract, how many years performance the team managers should pay attention, and it will be explain the variation of the expected salary as much as possible.

The Optimal Model for the 1st regression---which model we tend to use for the 1st regression

From the analysis we know that the salary may not be determined by 1 year performance. In order to find out how many years performance we should use, we may need to make a further test. First let us start to run the OLS regression for expected salary with "personal characteristics" and the "**performance in contract year**"

Source	SS	df	MS	Number of obs =	929
				F(19, 909) =	184.56
Model	21452.6605	19	1129.0874	Prob > F =	0.0000
Residual	5561.04085	909	6.11775671	R-squared =	0.7941
				Adj R-squared =	0.7898
Total	27013.7014	928	29.109592	Root MSE =	2.4734

salary	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
draft	-.6581973	.4922844	-1.34	0.182	-1.624343	.3079487
height	.1450573	.0424938	3.41	0.001	.06166	.2284545
weight	.0049226	.00562	0.88	0.381	-.0061071	.0159522
nation	.4003744	.2331165	1.72	0.086	-.0571347	.8578836
allstar	.2420402	.056323	4.30	0.000	.1315021	.3525784
age	-.2023275	.0264122	-7.66	0.000	-.2541634	-.1504916
hit2p	1.31393	.1579092	8.32	0.000	1.004021	1.623839
mis2p	-.3839367	.1365491	-2.81	0.005	-.6519248	-.1159486
hit3p	2.304839	.4452272	5.18	0.000	1.431046	3.178631
mis3p	-.7130531	.2894664	-2.46	0.014	-1.281153	-.1449531
hitFT	.8618201	.117111	7.36	0.000	.6319807	1.09166
misFT	.5444168	.2755649	1.98	0.048	.0035994	1.085234
ORB	.5552692	.1994409	2.78	0.005	.1638509	.9466874
DRB	.0759248	.1033366	0.73	0.463	-.1268812	.2787309
AST	.6690689	.0928513	7.21	0.000	.4868412	.8512967
STL	.500973	.304249	1.65	0.100	-.0961392	1.098085
BLK	1.406793	.2397938	5.87	0.000	.9361796	1.877407
TOV	-.5052848	.2880707	-1.75	0.080	-1.070646	.0600761
PF	-.5480261	.1771915	-3.09	0.002	-.895778	-.2002742
_cons	-8.519027	2.903769	-2.93	0.003	-14.2179	-2.820157

Table 3.1 The regression of salary and performance in contract year

Based on the result, we can easily find the coefficients and test statistics for each variable are reasonable and individually significant, when we just focus on the "performance in contract year".

However, we might think that the average salary of the contract might not just be determined by the performance in contract year, maybe the performance in earlier years would also influence the contract. Let us run an OLS regression for expected salary with the "personal characteristics " and "**the performance in both contract year and one year before contract year**"

Source	SS	df	MS	Number of obs =	894
Model	21320.5232	32	666.26635	F(32, 861) =	114.51
Residual	5009.83091	861	5.81861894	Prob > F =	0.0000
				R-squared =	0.8097
				Adj R-squared =	0.8027
Total	26330.3541	893	29.485279	Root MSE =	2.4122

salary	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
draft	-.6645934	.4886027	-1.36	0.174	-1.623585	.2943983
height	.1746715	.0438404	3.98	0.000	.088625	.2607181
weight	-.0004774	.0057121	-0.08	0.933	-.0116886	.0107338
nation	.3575416	.2335917	1.53	0.126	-.1009343	.8160174
allstar	.211671	.0582919	3.63	0.000	.0972602	.3260819
age	-.2419692	.0275622	-8.78	0.000	-.2960661	-.1878723
hit2p	1.198744	.1811182	6.62	0.000	.8432593	1.554229
mis2p	-.2816599	.1611833	-1.75	0.081	-.598018	.0346982
hit3p	2.41951	.4557901	5.31	0.000	1.52492	3.3141
mis3p	-1.074817	.3335057	-3.22	0.001	-1.729396	-.4202375
hitFT	.8102647	.1673113	4.84	0.000	.481879	1.13865
misFT	-.248051	.3762808	-0.66	0.510	-.986586	.490484
ORB	.2499959	.2446811	1.02	0.307	-.2302453	.7302372
DRB	.1338565	.1375948	0.97	0.331	-.1362039	.4039169
AST	.3354097	.1482071	2.26	0.024	.0445202	.6262992
STL	.6151625	.4092288	1.50	0.133	-.1880403	1.418365
BLK	1.40001	.3603128	3.89	0.000	.6928161	2.107204
TOV	-.1157794	.3344696	-0.35	0.729	-.7722506	.5406918
PF	-.7791714	.2273236	-3.43	0.001	-1.225345	-.3329981
hit2p_1	.4361808	.1937043	2.25	0.025	.0559929	.8163686
mis2p_1	-.4189332	.1636792	-2.56	0.011	-.7401902	-.0976762
hit3p_1	1.001335	.5292219	1.89	0.059	-.0373813	2.040051
mis3p_1	-.1789419	.3411061	-0.52	0.600	-.8484387	.490555
hitFT_1	.1509634	.1757775	0.86	0.391	-.1940391	.4959659
misFT_1	.7405376	.3859935	1.92	0.055	-.0170608	1.498136
ORB_1	.5435369	.2685044	2.02	0.043	.0165371	1.070537
DRB_1	-.1882171	.132538	-1.42	0.156	-.4483524	.0719182
AST_1	.3887565	.1471918	2.64	0.008	.0998598	.6776532
STL_1	.2141006	.4056417	0.53	0.598	-.5820617	1.010263
BLK_1	-.0073516	.3597195	-0.02	0.984	-.7133813	.698678
TOV_1	-.6967111	.3529314	-1.97	0.049	-1.389418	-.0040045
PF_1	.5828223	.2310125	2.52	0.012	.1294087	1.036236
_cons	-9.194039	3.064636	-3.00	0.003	-15.20907	-3.179007

Table 3.2 The regression of salary and performance in contract year and one year before that

After finding the performance before the contract year could influence the expected future salary, we try to add more previous year's performance and look at the result. Let us run a regression about salary with "the personal characteristics" and "**the performance in contract year, one year before contract year and two years before contract year**".

Source	SS	df	MS	Number of obs =	804
				F(45, 758) =	76.05
Model	19688.4469	45	437.521042	Prob > F =	0.0000
Residual	4361.01455	758	5.75331734	R-squared =	0.8187
				Adj R-squared =	0.8079
Total	24049.4614	803	29.9495161	Root MSE =	2.3986

salary	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
draft	-.7297473	.4971614	-1.47	0.143	-1.705724	.2462295
height	.1807	.0478368	3.78	0.000	.0867917	.2746083
weight	-.0009892	.0060908	-0.16	0.871	-.0129461	.0109676
nation	.2290048	.2485775	0.92	0.357	-.2589773	.716987
allstar	.2242152	.0613651	3.65	0.000	.1037496	.3446809
age	-.2560961	.0308962	-8.29	0.000	-.3167484	-.1954437
hit2p	1.111982	.193098	5.76	0.000	.7329111	1.491052
mis2p	-.230118	.1697549	-1.36	0.176	-.5633635	.1031276
hit3p	2.047846	.477334	4.29	0.000	1.110792	2.9849
mis3p	-.8856285	.3528073	-2.51	0.012	-1.578224	-.193033
hitFT	.8876489	.178368	4.98	0.000	.5374948	1.237803
misFT	-.3941839	.4144531	-0.95	0.342	-1.207796	.4194283
ORB	.1553826	.258789	0.60	0.548	-.3526456	.6634108
DRB	.2278767	.1475776	1.54	0.123	-.0618327	.5175862
AST	.279973	.1580507	1.77	0.077	-.0302961	.5902421
STL	.5178029	.442403	1.17	0.242	-.3506778	1.386284
BLK	.8381468	.3981465	2.11	0.036	.056546	1.619747
TOV	-.1536078	.3686307	-0.42	0.677	-.8772662	.5700506
PF	-.5366573	.2528972	-2.12	0.034	-1.03312	-.0401951
hit2p_1	.4267236	.2163608	1.97	0.049	.001986	.8514612
mis2p_1	-.2473724	.1930672	-1.28	0.200	-.6263824	.1316377
hit3p_1	1.328439	.5745579	2.31	0.021	.2005251	2.456353
mis3p_1	-.7448066	.3940907	-1.89	0.059	-1.518445	.0288322
hitFT_1	.0121355	.2044478	0.06	0.953	-.3892157	.4134867
misFT_1	.2017925	.4593686	0.44	0.661	-.6999932	1.103578
ORB_1	.5334124	.3178617	1.68	0.094	-.0905815	1.157406
DRB_1	-.0910696	.1581852	-0.58	0.565	-.4016026	.2194635
AST_1	.3119197	.1772808	1.76	0.079	-.0361	.6599395
STL_1	.1299989	.4811372	0.27	0.787	-.8145207	1.074519
BLK_1	-.5816766	.4434848	-1.31	0.190	-1.452281	.2889278
TOV_1	-.2010465	.3947592	-0.51	0.611	-.9759977	.5739048
PF_1	.5470528	.2801261	1.95	0.051	-.0028623	1.096968
hit2p_2	.0334832	.2116145	0.16	0.874	-.3819369	.4489033
mis2p_2	-.2242606	.1875096	-1.20	0.232	-.5923604	.1438392
hit3p_2	.8894421	.546481	1.63	0.104	-.1833539	1.962238
mis3p_2	-.0291954	.3619957	-0.08	0.936	-.7398288	.6814379
hitFT_2	.1205394	.1825145	0.66	0.509	-.2377546	.4788334
misFT_2	.8745439	.4142347	2.11	0.035	.0613602	1.687727
ORB_2	.1153077	.3081771	0.37	0.708	-.4896743	.7202897
DRB_2	-.2576985	.1505538	-1.71	0.087	-.5532504	.0378533
AST_2	.2722486	.1565772	1.74	0.082	-.0351279	.579625
BLK_2	1.288055	.4176342	3.08	0.002	.4681984	2.107912
STL_2	.251434	.4788762	0.53	0.600	-.6886471	1.191515
TOV_2	-.8418876	.3667727	-2.30	0.022	-1.561899	-.1218767
PF_2	-.0269262	.2546025	-0.11	0.916	-.5267359	.4728836
_cons	-9.307279	3.408885	-2.73	0.006	-15.99926	-2.615303

Table 3.3 The regression of salary and performance in contract year and 2 years before that

Unfortunately, most of the coefficients of previous years messed up and most of the test statistics are not significant any more. And the R square does not increase a lot which means the performance in 2 years before contract year cannot explain more variation of salary compared with other models. That result shows that we should not consider the performance 2 years before the contract year, it makes sense. When a team manager consider about the contract to provide, they must pay more attention to the most recent performance (the performance in contract year), maybe the less recent performance (the performance in 1 year before contract year), but generally they may not care about the earlier performance. It is reasonable, the player's performance might change a lot within a certain period, using too early performance as a judgment seems not wise.

As a final choice, we will use the 2nd model (table 3.2 in page 24), the model with "**the personal characteristics**" and "**the performance in contract year and one year before that**".

CHAPTER FOUR
Analysis for the 1st Regression

Source	SS	df	MS		
Model	21320.5232	32	666.26635	Number of obs	= 894
Residual	5009.83091	861	5.81861894	F(32, 861)	= 114.51
				Prob > F	= 0.0000
Total	26330.3541	893	29.485279	R-squared	= 0.8097
				Adj R-squared	= 0.8027
				Root MSE	= 2.4122

salary	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
draft	-.6645934	.4886027	-1.36	0.174	-1.623585	.2943983
height	.1746715	.0438404	3.98	0.000	.088625	.2607181
weight	-.0004774	.0057121	-0.08	0.933	-.0116886	.0107338
nation	.3575416	.2335917	1.53	0.126	-.1009343	.8160174
allstar	.211671	.0582919	3.63	0.000	.0972602	.3260819
age	-.2419692	.0275622	-8.78	0.000	-.2960661	-.1878723
hit2p	1.198744	.1811182	6.62	0.000	.8432593	1.554229
mis2p	-.2816599	.1611833	-1.75	0.081	-.598018	.0346982
hit3p	2.41951	.4557901	5.31	0.000	1.52492	3.3141
mis3p	-1.074817	.3335057	-3.22	0.001	-1.729396	-.4202375
hitFT	.8102647	.1673113	4.84	0.000	.481879	1.13865
misFT	-.248051	.3762808	-0.66	0.510	-.986586	.490484
ORB	.2499959	.2446811	1.02	0.307	-.2302453	.7302372
DRB	.1338565	.1375948	0.97	0.331	-.1362039	.4039169
AST	.3354097	.1482071	2.26	0.024	.0445202	.6262992
STL	.6151625	.4092288	1.50	0.133	-.1880403	1.418365
BLK	1.40001	.3603128	3.89	0.000	.6928161	2.107204
TOV	-.1157794	.3344696	-0.35	0.729	-.7722506	.5406918
PF	-.7791714	.2273236	-3.43	0.001	-1.225345	-.3329981
hit2p_1	.4361808	.1937043	2.25	0.025	.0559929	.8163686
mis2p_1	-.4189332	.1636792	-2.56	0.011	-.7401902	-.0976762
hit3p_1	1.001335	.5292219	1.89	0.059	-.0373813	2.040051
mis3p_1	-.1789419	.3411061	-0.52	0.600	-.8484387	.490555
hitFT_1	.1509634	.1757775	0.86	0.391	-.1940391	.4959659
misFT_1	.7405376	.3859935	1.92	0.055	-.0170608	1.498136
ORB_1	.5435369	.2685044	2.02	0.043	.0165371	1.070537
DRB_1	-.1882171	.132538	-1.42	0.156	-.4483524	.0719182
AST_1	.3887565	.1471918	2.64	0.008	.0998598	.6776532
STL_1	.2141006	.4056417	0.53	0.598	-.5820617	1.010263
BLK_1	-.0073516	.3597195	-0.02	0.984	-.7133813	.698678
TOV_1	-.6967111	.3529314	-1.97	0.049	-1.389418	-.0040045
PF_1	.5828223	.2310125	2.52	0.012	.1294087	1.036236
_cons	-9.194039	3.064636	-3.00	0.003	-15.20907	-3.179007

From this model we can find most of the coefficients and test statistics are reasonable and significant, implicating that we should not only consider about the performance in contract year, but also consider about the performance in one year before contract year.

The adjusted R square showed that about 80% variation of salary could be explained by the variation of independent variables. 80% is really large which can be treated as the evidence that our OLS regression fits the data pretty good.

1. Analysis for personal characteristics

(1) undrafted players tend to earn less

Draft is a dummy, it equals 1 if the player is undrafted player, 0 if drafted.

The explanation is the undrafted players' expected salary will be $0.66 million less than drafted players when we hold other factors constant. It makes sense. Because each year 60 new guys will enter the league through the NBA Draft, if a player is an undrafted player, that means he was not as good as others in the beginning. Although a few of them could become star players by chance, most of them could not survive in NBA. Their expected salary should be less than drafted players. It is not individually significant, not a really strong evidence, but it negative coefficient still shows the negative effect of "being an undrafted player".

(2) height is more important than weight in NBA

As most previous research analysis, height is a very crucial variable. The result of our regression also proved that. The player tend to earn $0.17 million more with one more inch, holding others constant. Consider about the magnitude of the coefficient, being a taller player really helps.

Compared with height, neither the coefficient nor the test statistic of weight is significant, implicating that weight is less important than height in NBA.

It is significant at 99% level. A pretty strong evidence to show that being taller will tend to have more future salary in NBA.

(3) International player salary premium seems exists in NBA.

Nation is a dummy. It equals 1 if the player is international player, 0 for the US players.

Foreign players expected salary will be $0.35 million higher than domestic players, holding other factors fixed.

As what I stated before, foreign players always represent the potential market, it means it could earn more money for the team because a lot of foreign people who will pay attention to their own country's player, maybe spend money for that. The international players might not be as good as domestic players, but they can earn a little bit more because the potential market behind their nationality.

It is not significant, but the sign of coefficient tells us the salary premium by nationality might exist.

(4)All Star players tend to earn more

Allstar is how many all star games the player has played until his contract year.

I have made a statement and introduction about All Star, it does not just mean the player is better than others, it also means the player is more popular than most other players. Popularity means ability to attract audiences and earn money for the team. That is how they get higher paid than normal players. It could be explained as the player's expected salary will increase $0.21 million when he has played all star game one more time before signing the new contract, fix other factors constant.

The coefficient is individually significant at 99% confidence level. It is really a strong evidence to show the premium salary for being an All Star player exists in NBA, and the more time the player could be an All Star, the more expected salary he could get.

(5) Expected salary decreases significantly with age growing

Age is the age of the player in his contract year. As what I expect, the negative effect of growing old is really large, the player's predicted salary will decrease $0.24 million when he grow 1 year older, keep other variables fixed. It makes sense, older players always means more

injuries, less potential, slower speed and less incentive to get progress. It is taking risk to give an old player a big contract. Their expected salary should be less and less when they become older and older. We cannot expect the old players can jump as high and run as fast as young boys. Time and tide wait for no man.

It is significant in 99% level, implicates the negative effect on the expected salary with age growing up cannot be ignored.

2. Analysis of scoring ability

contrat year	**coeff**	**t test**	**1 year before contract year**	**coeff**	**t test**
hit3p	**2.41**	(5.31)	**hit3p_1**	**1.00**	(1.89)
hit2p	**1.19**	(6.62)	**hit2p_1**	**0.43**	(2.25)
hitFT	**0.81**	(4.84)	**hitFT_1**	**0.15**	(0.86)

Hit2p means how many 2 point shot (per game) the player could hit in his contract year, it could be explained as the players predicted salary will be $1.19 million higher when he can hit one more 2 point shot in his contract year, fix other factors constant.

It is significant at 99% level.

similar, hit3p means how many 3 point shot (per game) he could hit in his contract year, which could be interpreted as the players expected salary will increase $2.41 million when he can hit one more 3 point shot in his contract year, when we hold other variables fixed. Someone may also think the coefficient is too large, when we look back to the sum of the variables. We can find that the mean of hit3p is just 0.66. Actually there are not so many good 3 point shooters in the league, and the coefficient makes sense when we consider about the value of 3 point shot.

It is significant at 99% level.

hitFT means how many free throw (per game) he could hit in his contract year. The interpretation for this should be the player's expected salary will increase $0.86 million when he could hit one more free throw in his contract year when we hold other factors constant.

It is significant at 99% confidence level.

[1] In the contract year, hitting a 3 point shot brings more future income than hitting a 2 point shot or hitting a free throw, and hitting a 2 point shot brings similar future income like hitting a free throw.

Now let us test the difference of the effect on salary between 3 point hit, 2 point hit and free throw hit **in contract year**

```
. test hit3p=hit2p

 ( 1)  - hit2p + hit3p = 0

       F(  1,   861) =    5.85
            Prob > F =    0.0158

. test hit3p=hitFT

 ( 1)  hit3p - hitFT = 0

       F(  1,   861) =   10.79
            Prob > F =    0.0011

. test hit2p=hitFT

 ( 1)  hit2p - hitFT = 0

       F(  1,   861) =    2.01
            Prob > F =    0.1568
```

The result showed that in the contract year, hitting a 3 point shot has totally different effect on future salary than hitting a 2 point shot or free throw. However, it also showed that compared with hitting a free throw, hitting a 2 point shot has no significant different effect on future income.

Why the 3 point shot is more valuable than 2 point shot?

2 point shot includes so many different ways, like fade-away shot, a hook shot, a double pump, a slam dunk or just a layup or even bank shot. We know some of them are hard to finish like hook shot and fade-away shot, some of them are much easier like layup and bank shot which everybody can do. The difference of difficulty exists between fade-away 2 point shot and layup,

but the similarities is either way can just get 2 point, no matter how hard the way is, even the player makes a 720 dunk, he will still get only 2 points, not 3 or 4 points.

3 point shot is different, it is far from basket, long distance and hard to hit. Each time the player hits a 3 pointer, he can make 3 points for his team. Without a doubt, 3 point shot has more value than 2 point shot.

[2] in the year before contract year, hitting a 3 point shot has no different effect with hitting a 2 point shot or hitting a free throw on the future income. Or we can say the 3 ways of scoring have no significant different effect on the future salaries.

Then let us test the difference of the effect on salary between 3 point hit, 2 point hit and free throw hit **in year before contract year**

```
. test hit3p_1=hit2p_1

 ( 1)  - hit2p_1 + hit3p_1 = 0

       F(  1,   861) =    0.99
            Prob > F =    0.3204

. test hit3p_1=hitFT_1

 ( 1)  hit3p_1 - hitFT_1 = 0

       F(  1,   861) =    2.20
            Prob > F =    0.1383

. test hit2p_1=hitFT_1

 ( 1)  hit2p_1 - hitFT_1 = 0

       F(  1,   861) =    1.00
            Prob > F =    0.3165

. test hit3p_1=hit2p_1=hitFT_1

 ( 1)  - hit2p_1 + hit3p_1 = 0
 ( 2)  hit3p_1 - hitFT_1 = 0

       F(  2,   861) =    1.30
            Prob > F =    0.2731
```

Interesting, in the year before contract year, hitting a 3 point shot, hitting a 2 point shot and hitting a free throw have no significant difference on the effect to the future salaries. We can treat as they have similar effect on the expected contract salary.

[3] the same type of performance in different years have different effect on the expected salary

```
. test hit2p=hit2p_1

 ( 1)  hit2p - hit2p_1 = 0

       F(  1,   861) =    5.60
            Prob > F =    0.0182

. test hit3p=hit3p_1

 ( 1)  hit3p - hit3p_1 = 0

       F(  1,   861) =    3.42
            Prob > F =    0.0646

. test hitFT=hitFT_1

 ( 1)  hitFT - hitFT_1 = 0

       F(  1,   861) =    4.30
            Prob > F =    0.0383
```

Based on the coefficients of them, we finally found that hitting a 3 point shot in contract year has more positive effect on the future salary than hitting a 2 point shot in the year before contract year. As the hitting 2 point shot and free throw showed the same.

From the table we can find that hitting one more two point shot in the contract year could bring **$1.19** million future expected income, but hitting one more two point shot in the year before contract year could only bring **$0.43** million to the predicted future income.

Similarly, hitting one more three point shot (**$2.41** million) in the player's contract year has more positive effect to the future predicted income than hitting one more three point shot(**$1.00** million) in the year before contract year.

The same situation exists in the hitting one more free throw (**$0.81** million) in the contract year and one year (**$0.15** million) before that.

The result shows that people always pay more attention to the most recent performance, especially to the performance in the contract year. For the same type of variable in different years,

the performance in the most recent year (contract year) always has more influence to the future salary than the same kind of performance in earlier years.

[4] The particularity of free throw

Based on the results we got above, hitting a 2 point shot and hitting a free throw in contract year do not have much different effect on the future salary.

However, we know that one free throw only worth 1 point, but 1 two point shot worth 2 points. How is that possible, does it mean 2 points equals 1 point?

Furthermore, if we look at the coefficients and test statistics carefully, we can find that the punishment on salary for missing a free throw shot is the least among 2 point miss, 3 point miss and free throw miss. The coefficient of misFT_1 (missing free throw in 1 year before contract year) is even positive and individually significant.

We noticed that, in the contract year, the negative effect of missing a free throw ($0.24m) is much less than the negative effect of missing a 2 point shot ($0.28m) or a missing a 3 point shot ($1.07m). At the same time, the coefficient of missing a free throw is not individually significant.

This makes sense if we think about free throw carefully. First let us think about 2 point and 3 point shot. When a player missed a 2 point shot, it just means he made a 2 point shot attempt and failed. If he is important enough to the team, he could make as many 2 point shot attempts as he wants and nobody dares to blame him because he is the star player to the team. Similarly, good shooter always makes a lot of 3 point shot when he is on fire, they can make unlimited time 3 point attempts if they really want.

However, the free throw is totally the different story. When the player misses a free throw shot, he must get the chance of free throw attempt first. How can a player get free throw? In most situation is because they draw a personal foul, then they can get the free throw chance, that is also explained why the free throw is called as “foul shot”.

What kind of player will always draw a foul? The center and power forward could always attack the basket easier because their position is closer to the basket, if they are skilled and strong enough, they can always draw a foul. Or maybe a speedy guard who is good at crossover and driving layup, could also have more chance to draw a foul. Anyway, missing a free throw means getting a chance of free throw, and that presents the ability of drawing a foul.

Drawing a personal foul does not just bring a free throw chance, it hurts the opponent even more. There is limitation for personal fouls in NBA. A player who is good at drawing a foul can hurt the opponent heavily because he can limit the performance of opponent. Because nobody wants to be kicked out of the basketball court for too many personal fouls. That is the extra value behind the free throw. It explained why hitting a free throw has the similar effect to future salary just like hitting a 2 point shot in the contract year.

3.Offensive rebound is important than Defensive rebound.

ORB means how many offensive rebounds (per game) the player could get in the contract year. That could be explained as when the player got one more offensive rebound in his contract year, his expected salary will increase $0.25 million, holding other factors fixed.

Similarly, DRB is how many defensive rebounds (per game) the player could get in the contract year. It leads to $0.13 million expected salary increase when the player could get one more defensive rebound in the contract year when we fixed other factors.

Neither of them is individually significant. Because our model includes 2 years performance, and I believe they are highly correlated with each other, which makes it insignificant, we test the jointly significance then.

```
. test ORB ORB_1

 ( 1)  ORB = 0
 ( 2)  ORB_1 = 0

       F(  2,   861) =    5.43
            Prob > F =    0.0045

. test DRB DRB_1

 ( 1)  DRB = 0
 ( 2)  DRB_1 = 0

       F(  2,   861) =    1.01
            Prob > F =    0.3636
```

We found evidence to show the coefficients of offensive rebounds for different years are jointly significant at 99% level.

Unfortunately, we do not find the evidence that defensive rebound is individually significant or jointly significant.

The offensive rebound of contract year and the year before that are jointly significant, but defensive rebound are not. That is an evidence that the value offensive rebound is much higher than the value of defensive rebound to their future expected salary.

“Good rebounders understand the game and personnel. They study who shoots, when and from where.”---- <How to Improve Basketball Rebounding: Tips, Drills and Fundamentals>

Technically, offensive and defensive rebounds are totally different things. The difficulty of getting an offensive rebound is much harder than getting a defensive rebound. Actually, people always say “grabbing” an offensive rebound and “protecting” a defensive rebound.

In basketball court, most of the rebounds are defensive rebound, the team on defense is much easier to get the missed shot because they can stay closer to the basket by rule, which means even a short point guard could get a lot of defensive rebound if they are smart enough and occupy a good position near the basket, like Jason Kidd(retired) and Rondo.

However, on the other hand, getting an offensive rebound is another story, they cannot stay closer than the team on defense by rule. When they want to grab an offensive rebound, they need

smarter brain, faster move and higher jump because they are always in the weak position compared with the team on defense.

Moreover, an offensive rebound means “another extra chance to attack” the opponent’s basket. That is an extra chance to make 2 point or 3 point shot or even draw a foul and make free throw, which is much more valuable and crucial than defensive rebound, obviously.

Let us look at the stats of ORB and DRB in contract year, and we can see the difference clearly. The defensive rebound is about 3 times larger than offensive rebound in average, implicating the difficulties of grabbing an offensive rebound.

Variable	Obs	Mean	Std. Dev.	Min	Max
ORB	929	1.142411	.8973434	0	5
DRB	929	3.093649	1.806493	0	10.7

As we know, a good offensive rebounder always needs strong muscle, keen sense of positioning and faster and higher jump, which leads to higher expected future salary.

The result also tells us compared with the ability of grabbing offensive rebound, the ability of protecting defensive rebound is less important, which might lead it (DRB) insignificant.

4.Analysis of Assist-----Assist is not as important as scoring point in contract year

AST is how many assists the player can make (per game) in his contract year. It means they player’s expected salary will increase $0.3 million due to one more assist he made in his contract year, Ceteris paribus.It is individually significant at 95% level.

As most people know, assist is passing the ball to the teammate who is in a good position to attack the basket directly. Not every good passing could be an assist, unless the teammate scores successfully. We can treat assist like the ability of “helping teammates get points”. In my view, assist is really important, even more important than scoring the points. Think about Steve Nash(retired 2014), who can contributed 15 points and 11 assists per game from 2004-2007, 11 assists

means 11 chances of getting 2 or 3 points. If 1 person could contribute about 40 points per game to the team, why do not the team managers pay higher salary to the player?

However, the coefficient (absolute value) of assist is less than the coefficients of hit2p, hit3p and hitFT, implicating that ability of “helping teammates get points” is less valuable than the ability of “getting points by myself” in the contract year. Let us make the test for that.

```
. test AST=hit2p

 ( 1)  - hit2p + AST = 0

       F(  1,   861) =   12.16
            Prob > F =    0.0005

. test AST=hit3p

 ( 1)  - hit3p + AST = 0

       F(  1,   861) =   19.12
            Prob > F =    0.0000

. test AST=hitFT

 ( 1)  - hitFT + AST = 0

       F(  1,   861) =    4.49
            Prob > F =    0.0344

. test AST=hit2p=hit3p=hitFT

 ( 1)  - hit2p + AST = 0
 ( 2)  - hit3p + AST = 0
 ( 3)  - hitFT + AST = 0

       F(  3,   861) =    9.25
            Prob > F =    0.0000
```

```
. test AST_1=hit2p_1

 ( 1)  - hit2p_1 + AST_1 = 0

       F(  1,   861) =    0.04
            Prob > F =    0.8482

. test AST_1=hit3p_1

 ( 1)  - hit3p_1 + AST_1 = 0

       F(  1,   861) =    1.22
            Prob > F =    0.2696

. test AST_1=hit2p_1

 ( 1)  - hit2p_1 + AST_1 = 0

       F(  1,   861) =    0.04
            Prob > F =    0.8482

. test AST_1=hitFT_1

 ( 1)  - hitFT_1 + AST_1 = 0

       F(  1,   861) =    1.03
            Prob > F =    0.3102

. test AST_1=hit3p_1=hit2p_1=hitFT_1

 ( 1)  - hit3p_1 + AST_1 = 0
 ( 2)  - hit2p_1 + AST_1 = 0
 ( 3)  - hitFT_1 + AST_1 = 0

       F(  3,   861) =    0.91
            Prob > F =    0.4379
```

Obviously, the assist worth less than hitting 2 point shot, 3 point shot or free throw in the contract year. However, in the year before contract year, making an assist has similar influence on future salaries just like hitting a 2 point, 3 point shot or hitting free throw. But the coefficients of assist are always less than the coefficients of hit2p, hit3p and hitFT in both years, which proved that the ability of “helping teammates get points” is less important on the future income than “getting points by myself”.

5.Steal and Block

(1) Why the Test Statistic of Steal is not significant

Steal is the defensive player cause a turnover of the opponent player by catching a pass or just taking the ball without touching opponent player's hand.

STL is how many steals the player could make in his contract year, explained like one more steal in the contract year can lead to $0.54 million expected salary increase, holding other factors equal. We can see the coefficient of STL is not individually significant, I would rather believe it is caused by multi-collinearity between steal in contract year and a year before that. Let us make a test.

```
. test STL STL_1

 ( 1)  STL = 0
 ( 2)  STL_1 = 0

       F(  2,   861) =    3.23
            Prob > F =    0.0399
```

They are jointly significant at 5% level.

Every time the player makes a steal, he decreases one opportunity of opponent to attack the basket and increases one chance of his team to attack opponent's basket. It sounds like really nice and it should be very significant at any level.

In my opinion, the reason that the steal is not so individually significant might be like this: a steal means a gamble, every time the defender tended to steal the ball, he lost his position, if he did it, it is fine; if he failed, he would have no time to recover in time. Moreover, we can think about who will always try to steal, without any doubt, the answer should be guard players. When a guard player loses his position, the opponent player will directly drive to the basket and the defender's gambling action will lead to the situation that 4 teammates defend 5 opponent players. It is taking risk.

Furthermore, in NBA level, nobody is easy to be stolen, every successful steal means more unsuccessful steal attempts because nobody could make it every time. However, consider about the benefit the steal could bring, increasing 1 more chance for the team to score and decrease 1 more chance of opponent to score, it sounds good. The positive coefficient shows that steal still has positive effect on the predicted future income, although it is not so certain as I thought.

(2)Why the coefficient of Block is so large

Block is the defensive player deflects the shot attempt of opponent player without touching opponent player's hand. It is how many blocks the player could make (per game) in his contract year. The interpretation of the coefficient of BLK could be the expected future salary will increase $1.40 million when the player could make one more block in his contract year, holding other factors fixed. It is individually significant at 99% level.

When we first see the coefficient, we might think it is a little bit large, ever larger than the effect of hitting one more 2 point shot. So let us look at the mean and standard deviation of block.

Variable	Obs	Mean	Std. Dev.	Min	Max
BLK	929	.4966631	.5241985	0	3.7

We can clearly see that the mean is 0.49, which means the players could not make even half block per game in average in their contract year. Even the best blocker could only make 3.7 blocks per game. Consider my database includes performance in 15 years, we can understand how difficult blocking a shot would be.

Why block is so difficult? Blocking a shot is also a kind of gambling action, if the player did it, he made a successful defense and stopped opponents shot. However, If he failed, he would lose his defensive position because he always need to jump to make block attempt, even worse, he might increase a personal foul to himself if he touched the opponent player's hand, we know that a player will be kicked out of playing court if he made more than 5 personal fouls. At the same

time, he would send opponent player to the free throw line, consider about the percentage of free throw in NBA, making an unsuccessful block attempt almost equals giving opponent 2 point for free.

Let us think about what kind of player can always have chance to make a block. The answer should be center and power forward, because their position is closer to the basket and generally they are tall and strong enough. The center and power forward are always considered as the major defenders of the team and they always need to face the opponent guard players who always drive to the basket. The guard players are always faster and more explosive, when they make a fast break and charge toward the basket, it is hard to block because trying to make a block will increase the risk to get a personal foul---- it is easy to touch opponent player's hand or body when they charge close to the basket with high speed. The opponent players could drive inside as many times as they want, but each defensive player can only have at most 5 chances to make personal fouls.

That is also explained why one more block in contract year can bring so much expected future income. Being a good blocker does not just mean deflecting a shot, it is also a warning to opponent players. The opponent players will tend to make less driving layup because they know they might be blocked. Being a good blocker can decrease opponent players driving layup attempts, this effect cannot be showed in stats of performance, but it exists, and it should be highly correlated with block, and it is how they got paid by having this ability.

6.offensive performance seems like more important than defensive performance in the year before contract year

Unlike the offensive rebound and assist, neither of the steal and block in the year before contract year is significant. That could be an evidence to show that people tend to pay less attention to that kind of early performance. After all, nobody will really care about how many

steal and block the player could make 2 years ago, they might good at stealing the ball, but it is far from now, no one can make sure they still have the ability that he had before 2 years. Relatively, paying more attention to what they could do 1 year ago is more sensible, that is why the coefficients of steal and block are significant in contract year. Furthermore, even though the team manager believes the player is still good at stealing or blocking, how much do the stealing and blocking worth, compared with the ability to score (2 point, 3 point shot)? Anyway, I have never heard the audience will buy ticket to watch stealing and blocking show, and the result of the model has proved what I thought.

From what we analyzed before, the ability of getting points (hit2p, hit3p and hitFT in both years) are always more important than the ability of helping teammates get points (assist). At the same time, offensive rebound is always more valuable than defensive rebound.

Then we can conclude that scoring points(2p, 3p and free throw), helping teammates score points(assist) and offensive rebound are the most important performance among all the performance.

And if we think about it carefully, all the 3 different types of performance (getting point directly, getting another chance to score, passing the ball to teammate to score) are just one thing----scoring the points. Actually, players seem like got paid in the future by their ability of ***"helping the team to score points"*** in the contract year and one year before that. That could explain why the coefficients of defensive rebound and steal or block are not as significant as the coefficients related to scoring points. And I call that performance as "**offensive performance**".

If we think about the on-court performance carefully, we can roughly divide it into 3 parts. **(1)Offensive performance**, which means the performance that can help team score more points directly or indirectly. Like assist, 2 point shot, 3 point shot, free throw and offensive rebound. **(2)Defensive performance**, which means the performance that can help team lose less points

directly or indirectly. Like block, steal and defensive rebound.

(3)Negative performance, it only hurts, never help. Just like personal foul and turnover.

According to the result of our regression, the coefficients of offensive performances are always jointly significant or individually significant in both years. However, the defensive performance and negative performance are not always significant, especially in the year before contract year. More interesting, the magnitudes of the coefficients of offensive performance are always larger than defensive performance in both years, represents offensive performance has larger effect to the expected salary. This is pretty funny, implicates that the team managers value the ability of "**helping team score points**" more than any other performance when they offer a contract.

Why the expected future salary depends heavily on the ability of "**helping team score points**"? In my opinion, basketball is different with soccer or any other kind of sports like what we analyzed in the beginning. We cannot expect to win if the team only gets 10 points per game. More points always bring higher chance to win, winning the game can create a lot of stable fans, it might help the team increase tickets sale and licensed good sale. Furthermore, the tickets price of playoff games will be much higher than the regular games, and more people will tend to watch the playoff because it is related to the final champion. Getting more points can win the game, winning the game can create box office revenue, which will increase the team's income.

Based on my regression, the performance of defense are less important than the performance of offense in the year before contract year. However, it is not a solid evidence to make the conclusion that "offense is better than defense (for players' future income)".

We can think about our measurement of offensive performance. 2 point shot, 3 point shot, free throw, assist and offensive rebound. It almost includes everything related with offense. But our defensive performance measurement only includes steal, block and defensive rebound. It

seems nice at first glance because a lot people believe that is what defense means on the playing court. Sure, it is defense, but it does not include all the defensive action.

The measurement of defensive performance is much more complicated than the measurement of defensive performance. Without a doubt, steal, block and defensive rebound is the performance that can help team lose less points. However, there are a lot of unmeasured factors which can influence the defense quality. For example, the sense of positioning or the ability of block out can help teammate get the defensive rebound, however, it will not showed in the player's stats because he is not the person who protects the rebound. However, without his help the teammate could not get the rebound. Furthermore, let us consider about the type of defense in playing court. The most common defense is man-to-man defense which relies on the single player, however, there are also zone-defense and combination defense which rely on the team. In most situation, the defense quality depends on the whole team, each player on the court plays the right role and appears in the right position. A successful defense relies on a lot of players, but not just one player, even a block or a steal is always correlated with teammates help, although it will not be showed in their stats.

Based on the difficulties of measuring the performance of defense, I suppose the result we found (offense is more important than defense to the players' future income) might be not as solid as I thought. It might be a weak evidence that "helping team get more points" is more crucial than "helping team lose less points", or it might be wrong. At least, we can still find that compared with steal, block and defensive rebound, the ability of getting points is more important in the year before the contract year.

CHAPTER FIVE

Analysis for the 2nd Regression

Introduction of the variables in the 2nd regression

Dependent variable	
Diff	the difference between the predicted salary and the real average salary of the new contract in the year of signing the new contract, use million dollar as unit.
Independent variables	
Length	the length of the contract, use year as unit. If the contract is 5 years, $80 million in total, the length =5
Sum	the sum amount of the contract, use million dollar as unit. If the contract is 5 years, $80 million in total, sum=80
Allstar	how many All Star games the player has played until his contract year
Age	how old the player was in his contract year
PTS	how many points(per game) the player could get in the contract year
TRB	how many rebounds(per game, offensive plus defensive) the player could get in the contract year
AST	how many assists (per game) the player could get in the contract year
BLK	how many blocks (per game) the player could make in the contract year
STL	how many steals (per game) the player could make in the contract year
TOV	how many turnovers (per game) the player made in the contract year
PF	how many personal foul (per game) the player made in the contract year

Summary of the data

Variable	Obs	Mean	Std. Dev.	Min	Max
diff	809	-.2590546	2.770091	-15.0181	8.304125
length	955	3.008377	1.779199	1	7
sum	955	22.46686	26.62812	.46	136.4
age	835	27.30659	3.514873	20	39
allstar	955	.5151832	1.756846	0	15
PTS	835	10.00228	6.181851	0	28.9
TRB	835	4.111257	2.572408	0	16.3
AST	835	2.173892	1.998258	0	11.5
STL	835	.748982	.4592072	0	2.8
BLK	835	.4929341	.5334476	0	3.3
TOV	835	1.409581	.8493215	0	5.7
PF	835	2.088263	.7311049	0	4.3

Tabel 5.1 summary of the data for 2nd regression

Empirical Model for the 2nd Regression

After getting the coefficients of determinants of salary, we can use these coefficients to get the expected future salary in the year of signing a new contract, then get the difference between the predicted salary and real salary. By analyzing the difference, we would see which type of contract and what kind of players tend to be overpaid or underpaid.

The 2nd model will be like following:

Diff=β0+β1*contract length+β2*contract sum+β3*age+β4*allstar+

+β5*performance in contract year+ε

Where the **diff** represents the difference between “the predicted salary” and “the real average salary of the new contract” in the year of signing the contract.

The 2nd model will tell us which type of contract and what kind of players tend to be overpaid in the year of signing a new contract. As the result, the 2nd model will be a valuable reference for both the team managers and players. It will let the players know what kind of performance is more important to their future income, and it will also tell the boss and managers that what type of contract they should provide to maximize their profit.

When I got all the coefficients which decide the expected salary from the 1st regression, I want to make a prediction to their future income in the year signed a new contract based on the performance and all the other variables.

For example, suppose a player signed a 5 years $30 million contract before 2006, his salary is decided by the performance in 2005 and 2004 in my first regression above. After getting the coefficients (which I have already had), I want to use the coefficients and the performance in 2006 and 2005 to get the predicted value of his expected salary in 2006. After getting that predicted value, I can use the predicted value to compare with the average salary ($6 million=30/5) of the real contract, and use the predicted salary of 2006 to minus real salary in 2006 and get the difference.

The differences mean some players are overpaid, while others are underpaid. I want to use the difference like the dependent variable, and use contract years performance and contract sum and contract length as independent variables and some other variables to analyze which kind of player and which kind of contract will lead to the player overpaid or underpaid in the year he signed the contract.

Why do I want to look at the performance and salary in the year he signed the contract? I always noticed one phenomenon, when the player signed a new contract, the new contract will always start several years later, like Melo, he signed a contract before season 2006, but the new contract will start from 2007. More surprisingly, although Melo would still get the low level salary in his rookie contract in 2006(the new contract works since 2007), his performance still increased a lot in 2006. I suppose it is because he got incentive by signing a fat contract. As we analyzed in the beginning, the guaranteed and multi-year contract made him knew exactly he would get it later although he could not get it immediately in 2006.

I believe signing a new contract will have significant influence to the player's performance. It could be a positive or negative incentive. Although my research is not just about player's performance, but the predicted salary is based on the player's performance in my previous model which makes me can get the predicted salary based on his current performance. Then I can look at the underpaid and overpaid issue.

Let us look at an example, Melo, he signed a 5 years $80 million contract before 2006 season started, the contract would start at 2007 which means Melo could only get $4.6 million in 2006 based on his rookie contract. However, does it really means his performance just worth $4.6 million? No. He signed a new contract with average $16 million per year is because the team managers believed Melo's performance was already worth $16 million per year, it is not only an investment for the future, but also a reward to his past. Moreover, it is also an incentive to his performance in current season. That is why I choose to compare the predicted salary with the contract average salary even if the contract might not start yet. Suppose I get the predicted salary of Melo in 2006 is 17 million, what is the difference? 17 minus 4.6 or 17 minus 16? Without a doubt, the last one, the difference will be 17-16=1.

The result of the 2nd regression

Source	SS	df	MS	Number of obs =	809
Model	4130.19907	11	375.472643	F(11, 797) =	144.57
Residual	2069.90977	797	2.59712644	Prob > F =	0.0000
				R-squared =	0.6661
Total	6200.10885	808	7.67340204	Adj R-squared =	0.6615
				Root MSE =	1.6116

diff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
length	.4026009	.0586041	6.87	0.000	.2875642	.5176375
sum	-.1501743	.0050088	-29.98	0.000	-.1600064	-.1403422
age	-.2384185	.0193266	-12.34	0.000	-.2763556	-.2004813
allstar	.1167372	.0389143	3.00	0.003	.0403506	.1931237
PTS	.392446	.0210744	18.62	0.000	.3510781	.4338139
TRB	.169935	.0408773	4.16	0.000	.0896951	.250175
AST	.2301775	.053846	4.27	0.000	.1244808	.3358743
STL	.3419755	.1971111	1.73	0.083	-.0449427	.7288937
BLK	1.320659	.1559493	8.47	0.000	1.014539	1.626779
TOV	-.6338187	.1590911	-3.98	0.000	-.9461058	-.3215315
PF	-.1598906	.1202816	-1.33	0.184	-.3959967	.0762156
_cons	3.649791	.6194539	5.89	0.000	2.433837	4.865745

Table 5.2 result of 2nd regression

The model could explain about 66% of the variation of the difference between the predicted salary and the real salary by all the independent variables.

Because I use predicted salary to minus the average salary of the contract, which means if the coefficient is positive, we underpaid the player, it is the money we saved in the year of signing the new contract, if the coefficient is negative, we overpaid the player, it is the money we wasted in the year of signing a new contract.

1.Contract length

Length means the length of the player's new contract. One more year increase of contract length can save $0.40 million for the team, holding other factors constant. It is significant at 99% level. When we hold the sum of the contract, a longer contract always means less payment per year. That could be an explanation why the players tend to be underpaid with a long contract.

2.Sum of the contract

Sum means the sum amount of the contract, which could be explained as 1 more million dollar increase of the sum amount of the new contract will lead to the team overpay the player $0.15 million in the year he signed the contract. It is significant at 99% level.

The coefficient of sum makes sense. We all know that fat contract may lead the players perform not as hard as before, because the player knows he will get a large amount of money in the future and most of them are guaranteed. They will vary their behavior to maximize their own utility just like decrease his effort to avoid of injury in order to sign another fat contract. When we look at the magnitude of the coefficient, 0.15, which can be treated as the team might overpay the player 15% of total contract salary in the year of signing a new contract. It is risky, however, because there are a lot of factors can influence the player's salary and contract, like how popular the player is and the ability to attract advertising and audience. We cannot measure and calculate that kind of factors, and the 15% might be just caused by that kind of reasons. I would rather believe that the team managers never overpay the players in average in the real world.

3.Age

Age is the age in the contract year. The coefficient could be interpreted as the team will overpay $0.23 million to the player in the year of signing a new contract if the player is one year older in his contract year, fixing other factors constant. It is significant at 99% confidence level.

Old players always performed not as good as before, that is true. However, we should think about it carefully. If old players cannot perform as well as before is a common sense, the team managers and the bosses should know it clearly than anybody else, because they pay the salary, not us. The team managers, are they not as smart as normal people? No, they should be more rational than others because they will lose hundred million bucks if they made a big mistake about the contract.

The result showed that signing an old player has risk of overpaying him. It might be true, but I suppose the story should be more like this: old players cannot run as fast and jump as high as before, it is because their physical condition decreases with age growing, that is true. However, old players also have more experience than the young players, they are more familiar with the rule and have more useful skills in playing court like “flop”, which means acting as he was fouled, or just some little tricks to protect their bodies from getting injured by opponents. Most of these abilities cannot be shown by performance stats, but it is really useful. Consider about Battier (retired 2014), his stats is nothing, but every team manager and coach likes him very much because his value is behind his stats.

Moreover, when the old player went to a young team with a lot of young boys, he could be a teacher to his young teammates, sharing his experience from his long career to the rookies can help them improve faster. Think about Andre Miller, an old point guard with plenty of experience, he was traded to Washington in 2013. The leader of Washington in 2013 was John Wall, a talented point guard who was born in 1990. John Wall admitted that he really learned a lot from Miller, a lot of little tricks and skills of being a good point guard. And next season in 2014, John Wall became a All Star player for the first time and lead Washington to the playoff and beat down Chicago bulls in 2014 playoff.

Similar instances are countless. An old guy could contribute more than the stats of the on-court performance, that is why I said the team managers never overpay the old players in average, it is just because some factors we cannot observed and measured, but the factors really influence his expected income and make the team manager give a contract seems like overpaid to the old guys. I am not saying nobody is overpaid, I have to admit that there are some players are overpaid based on every evidence, but there are also some players are underpaid, I would rather believe the

team never overpay or underpay the players in average, although offering an old guy a big contract is taking risk, as the result showed.

4.All Star

Allstar is how many times the player has played all star games until his contract year. And the coefficient could be explained like the team will underpay (save) $0.11 million in the year signing contract when the player has played the All Star games for one more time until his contract year, holding other factors fixed.

I like this result, the all star players are someone who have proved themselves, they are good at attracting audiences or they are top players, or even both. They have proved what they could do and built their own reputation in the league. Generally they will not decrease their effort once signing a fat contract. That is how they became All Star. And the more times he played All Star games, the lower risk to give him a contract because he will always work hard for their reputation and for the next contract. The result showed that signing an All Star player is a wise choice, at least it brings benefit to the team in the year of signing contract.

5.On-court performance

The points, rebound, assists, blocks, each of the coefficients is still individually significant at 90% level, which means with the increase of these types of performance in contract year, the team can underpay (save) the players in the year of signing a new contract.

For example, the team would underpay (save) $0.39 million to the player in the year he signing the contract if he could get one more point in his contract year, holding other factors constant.

All of the coefficients and the test statistics are reasonable and significant. For most of the performance like how many points or how many rebounds the player could get in contract year,

the better they performed in contract year, the more money they could save for the team in the year of signing a new contract.

We noticed that the only insignificant coefficient of on-court performance is personal foul. This problem may be caused by sample size. We will talk about it later in the last part of the paper.

Conclusion of 2nd regression

In conclusion, the 2nd model tells us in order to save money for the team in the year of signing a player, the optimal choice for the team is signing some young All Star players who are good at scoring points with a relatively small contract. I could remember most teams did the same thing over and over since I watched NBA. It makes sense.

The model also tells us signing following players has risk, like signing an old player, signing a player who could not get a lot of points or who made too many turnovers in his contract year or signing a player with a relatively big contract. In order to lower the risk, the team managers need to think carefully about which contract they should provide and which type of players they should sign.

CHAPTER SIX

The Weakness of the Paper

1.Omitted variable bias

Actually, there are a lot of factors outside the playing court could influence the contract the player could get. The most significant factor is price competition. Sometimes several teams want to get 1 player, the player's salary will be higher than the expected salary he deserved. Similar, in some special situation, some teams have no enough salary room because of the salary cap. Then some players will have to sign an underpaid contract. Or some performance which cannot be measured by stats like keen of positioning and ability of zone-defense, but it is correlated with the players' performance. This is not what we can control, but it really happened and makes our result biased somehow.

2.multi-collinearity

The other problem is high correlation between different variables because our model includes a lot of same type of performance in different years.

Like the rebounds per game for the same player in different seasons are highly correlated because the ability and skill of a certain player should be stable, generally it will not change too much in short period. But it will also influence the result of the regression and make the estimate insignificant.

3.Sample size

The 3rd problem is the sample size is not large enough, I found 955 samples of different 432 players within the past 15 years. However, I only got 894 samples for the first regression and 809 samples for the second. Each dependent variable is connected with multi-year performance, that means if one of the performances is missing value (injury or absence for any reason), I will lose this sample when I run the regression.

Based on the principle between new contract and performance, each player only provides 1 sample once he signed a new contract, so the sample for each player should equal N-1, (N is the number of contracts in whole career, 1 is the rookie contract). For instance, Yao Ming, a star player who played 8 seasons in NBA, only provides 1 sample in my database. We should notice this is not a specific example, based on the character of multi-year contract, a lot of players only provides 1 or 2 samples, even for Kobe, the super star who has played 18 seasons, only provides 4 samples. Not mention that a lot of players could not even get his 2^{nd} contract in NBA after the rookie contract, which makes them cannot be a sample of my data.

However, these weaknesses do not mean that my regression is useless although the coefficients are somehow biased. Based on the OLS assumption and properties, almost all the regression will have weakness, it always has omitted bias problem. However, we can still get some useful experience from this. Like how many years performance will influence the future salary and how to build a connection between performance and salary, and use the expected salary to compare with the real salary to find which type of player tend to be overpaid or underpaid in the year of signing contract. We finally found that the offensive performance is much more valuable than defensive and negative performance for searching a big contract in the future. That is really valuable for the team managers and players. It could tell the players which kind of performance worth more, and tell the managers which kind of contract they tend to use.

Reference

Jean, Neal. (2010)Performance variation in NBA---guaranteed contract and contract year phenomenon (working paper)

Stiroh, Kevin. (2007)Playing for keeps: Pay and performance in NBA

Economic Inquiry Volume 45, Issue 1, pages 145–161,January 2007

David J. Berri and Anthony C. Krautmann (2006) Shirking on the court: testing for the incentive effect of guaranteed pay. Economic inquiry, ISSN 0095-2583,
2006, Volume 44, Issue 3, pp. 536 – 546

Erick Eschker, Stephen J. Perez & Mark V. Siegler (2004) The NBA and the influx of international basketball players. Applied Economics, Volume 36, Issue 10, pp. 1009 - 1020

Wallace,Michael (1988) Labor market structure and salary determination among professional basketball players. work and Occupations August 1988 vol. 15 no. 3 page 294-312

Prendergast, Canice.(2008) "Intrinsic Motivation and Incentives." American Economic Review 98, no. 2 (2008): 201-205.

Wen-Jhan Jane, Min hua kuo and Sheng Tung Chen (2013)Does salary discrimination by nationality exist? In case of Nippon professional Baseball League Modern Economy, 2013, 4, 93-101

Hayles, Jame.(2006) Does an All Star Premium exist in NBA? An econometric analysis of NBA player salaries from 1999-2006 (working paper)

Joshi, Nikhil.(2011)The market for talent: Are top draft picks overpaid? (working paper)

How to Improve Basketball Rebounding: Tips, Drills and Fundamentals---breakthrough basketball.com

Alchian, Armen A., and Harold Demsetz. "Production , Information Costs, and Economic

Organization." American Economic Review 62, no. 5 (1972): 777-95.

Lehn, Kenneth (1982)Property Rights, Risk Sharing, and Player Disability in Major League Baseball. Journal of Law and Economics, Vol. 25, No. 2 (Oct., 1982), pp. 343-366

Fans ought to walk out on overpaid greedy NBA(1999). Spokesman Review

www.basketball- reference.com

www.spotrac.com/nba/ basketball.realgm.com

www.usatoday.com