STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic Year 2016-17

Do athletes' individual salaries depend on the performance of their peers?

Prototype heuristic in wage bargaining in the NBA

Marian Schwinner (40901)

Abstract

This thesis analyzes the link between relative market value of representative subsets of athletes in the National Basketball Association (NBA) and individual wages. NBA athletes are categorized with respect to multiple performance characteristics utilizing the k-means algorithm to cluster observations and a group's market value is calculated by averaging real annual salaries. Employing fixed effects and GMM estimation techniques, I find a statistically significant, positive effect of one-period lagged relative market value of an athlete's representative cluster on individual wages after controlling for past performance. This finding is consistent with the theory of prototype heuristic, introduced by Kahneman and Frederick (2002), that NBA teams' judgment about an athlete's future performance is based on a comparison of the athlete to the average properties of the representative group of similar individuals in that they offer a comparable role on the basketball court.

Keywords: Prototype heuristic, Wage bargaining, NBA, Behavioral Economics of Organization, Prototype wage
JEL: D82, J31, J44

Supervisor: Martina Björkman Nyqvist
Date submitted: January 2, 2017
Date examined: January 10, 2017
Examiner: Maria Perrotta Berlin
Discussants: Alexander Fält and Nga Nguyen

Acknowledgements

I am grateful for everybody who supported me in my research leading by my family. A special thanks to my supervisor, Assistant Professor Martina Björkman Nyqvist, to Professor Harald Oberhofer from the Vienna University of Business and Economics and to all members of the Austrian Institute of Economic Research who provided help and advice. Furthermore I would like to thank Professor Anna Dreber Almenberg and Maria Perrotta Berlin for help-full discussions and comments.

Contents

List of Figures

List of Tables

1 Introduction

One of the main results of agency theory is that properly designed contracts may align the interests of agents and the respective principal, if the contract provides incentives for the agent to choose the level of effort necessary to produce optimal output for the principal. The construction of such a contract is a rather easy computational task if the agent's actions are common knowledge to both parties. However, in most principal-agent settings it is either impossible or very costly for the principal to observe the agent's actions, which places special demand on the design of contracts. As proposed by Harris and Holmström (1982), past output may be taken as a proxy for the agent's willingness to show effort and her ability and, hence, form the basis for negotiations. The drawback is that output is often influenced by other variables than internal factors of the agent and the principal has to decide on the appropriate compensation under undesirable uncertainty.

The imperfect correlation of the agent's input and performance allows room for behavioral influences that are not considered in traditional agency theory. That is, the principal may be systematically biased in judging the cause of performance and ,therefore, overestimate the importance of effort and ability. Thus, success is often wrongly attributed to internal factors of the agent. Bertrand and Mullainathan (2001) show that compensation of oil company executives is positively linked to a raise in oil prices even though worldwide prices are set by global demand and the Organization of the Petroleum Exporting Countries (OPEC). Moreover, the executives' compensation is independent of a fall in prices. Even further, wage preferences may depend on reference points such as past compensation. Camerer, Babcock, Loewenstein, and Thaler (1997) show that cab drivers adjust their daily labor supply according to an income target and Fehr and Goette (2007) find similar results in a field experiment with bicycle messengers.

My research focuses on the principal's judgment of the agent's future performance and possible biases due to the fundamental uncertainty of the process. Tversky and Kahneman (1975) identified heuristics that may affect the principal's estimates of ability and future effort of the agent according to the level of representativeness with respect to situations that come easily to mind. Following this theory, the principal may base her judgment about the value of the agent on past situations that serve as a comparison to current negotiations. As a consequence, the agreed compensation may depend on a set of available observations that show enough similarity. Moreover, this research will concentrate on the analysis of *prototype heuristic* introduced by Kahneman and Frederick (2002), which states that judgments under uncertainty may be partly based on the principal's perception of the agent's prototype, the average value of salient properties of the "homogeneous" set the agent belongs to.

This thesis employs a unique data-set for professional basketball players in the National Basketball Association (NBA) from 2009 to 2016. The labor market in the NBA provides an ideal environment to study behavioral effects in contract negotiations due to rich data on athlete performance, individual characteristics and compensation structure. Moreover, the game of basketball is governed by a well defined set of rules and marginal contributions to team-success may be easily identified.

Athlete performance in the NBA is multi-dimensional and every player on the floor needs to fulfill a certain role for the team to be successful. In the framework of my research I will categorize athletes with respect to productivity dimensions in the hope to form representative sets for such athlete roles. Furthermore, prototype wages, average group-specific annual income, is calculated and may be interpreted as the relative market value of the more or less homogeneous group. The main hypothesis of this thesis is that NBA teams are influenced by past market value of the role of the athlete they are negotiating with, thus, over- or underestimating the individual value according to the past performance of athletes with a similar role.

The data provide broad support for this thesis' main hypothesis as one-period lagged role-specific market value does in fact positively affect individual salaries after controlling for individual performance. Given the visibility and competitiveness in professional sports, it is likely that NBA teams have a good basis for evaluating an athlete's ability and willingness to show effort by observing his on-court performance. Still, individual wages are additionally adjusted for past performance of the cluster the athlete is a member of, thus, clearly contradicting a main hypothesis of the simple principal-agent model that compensation is solely based on the agent's signaling of ability and effort.

In Section 2, I will motivate the empirical work with a theoretical agency model and discuss implications of judgment bias regarding the principal's payment decision under uncertainty. Then, in Section 3, I will give a brief background regarding the NBA, the institutions governing individual contracts and market entry, and how NBA athletes may be categorized by potential employers. Section 4 presents the utilized data, econometric methods and cluster-techniques. In Section 5, I will report estimation results considering different cluster-specifications and the robustness of the findings will be analyzed in Section 6. Section 7 will then finish by summarizing the empirical findings and discuss possible limitations of this thesis.

2 Theoretical framework and relevant literature

A large body of theoretical literature on optimal contract theory discusses agency costs and the conceptional problem of moral hazard in contract negotiations.[1] In this section, I will present a theoretical model of wage dynamics based on Harris and Holmström (1982) that illustrates the principal's payment decision under uncertainty. In the National Basketball Association (NBA) athletes are free to negotiate new contracts with any of the thirty teams. Also, they are allowed to negotiate contract extensions with current employers. Assuming that the athlete's effort is independent of the principal, all hypothetical future employers face the same judgment decision upon the athlete's ability that can only be estimated by observing past output, i.e. the athlete's performance in past seasons. This process under imperfect information may be prone to heuristics and therefore dependent on additional factors that may influence the estimation process of the agent's ability. A number of articles discuss the influence of judgment heuristics on market settings. For example, Kliger and Kudryavtsev (2010) find that daily market returns affect investor's reaction to analyst recommendations, Barber and Odean (2008) discuss the effect of recent news on buy-decisions of investors and Lee, O'Brien, and Sivaramakrishnan (2008) find that analysts significantly overweight recent market developments regarding long-term forecasts.

2.1 Principal-Agent model with assymetric information

I will consider a classic agency model in which an employee (agent) is compensated for costly effort by an employer (principal), whose success depends on the agent's effort and ability. The optimal contract maximizes the agent's utility from effort compensation while giving incentive to choose the level of effort that maximizes the principal's utility from production.

The moral hazard problem in a principal-agent setting arises when "full observation of actions is either impossible or prohibitively costly". (Holmström, 1979, p.74) As a consequence, both agent and principal are imperfectly informed about the agent's ability, which is only indirectly observed by production, i.e. the principal is not able to distinguish between true ability, effort and noise. Harris and Holmström (1982) proposed a model of wage dynamics in which both parties learn gradually about the agent's ability by observing output over time. Due to the imperfect correlation of production and ability, output can only serve as a proxy in a contract between principal and agent. Formally, an agent with ability $\eta_{i,t}$ and effort $e_{i,t}$

[1] See Grossman and Hart (1983), Hart and Holmström (1986) and Holmström (1979)

in period t will produce an output

$$y_{i,t} = f(\eta_{i,t}, e_{i,t}, \epsilon_{i,t}), \tag{1}$$

where $\epsilon_{i,t}$ is an error term which represents the fact that output is driven by other factors than ability and effort. The time subscript on ability reflects a dynamic development of agent i's ability and assures that agent i's effort is not converging to 0 if analyzed in a career setting. (Holmström, 1999)

Agent i has a Bernoulli Utility function over effort, e_i and compensation, W[2],

$$U_i(W_t, e_{i,t}), \tag{2}$$

which is assumed to be increasing in wage, $\frac{\partial U_i}{\partial W_t} > 0$, and decreasing in effort, $\frac{\partial U_i}{\partial e_{i,t}} < 0$. The impact of present effort on future wages determines the agent's effort decision.
Principal j's Bernoulli Utility function over income is defined as

$$U_j(y_{i,t}, W_t), \tag{3}$$

which is assumed to be increasing in output, $\frac{\partial U_j}{\partial y_{i,t}} > 0$, and decreasing in effort compensation, $\frac{\partial U_j}{\partial W_t} < 0$.

The history of outputs up to period t, $y^{t-1} = (y_1, ..., y_{t-1})$, is assumed to be common knowledge and forms the basis for effort compensation. Further, given the agent exceeds her reservation utility $\bar{U}$, principal j faces the following infinite horizon[3] utility maximization problem:

$$\max_{\boldsymbol{W}} \mathbf{E}[\sum_{t=1}^{\infty} \beta^{1-t} \cdot U_j(y_{i,t}, W_t)|y^{t-1}] \tag{4}$$

subject to the constraints

$$\mathbf{E}[\sum_{t=1}^{\infty} \beta^{1-t} \cdot U_i(W_t, e_{i,t})|y^{t-1}] \geq \bar{U} \tag{5}$$

[2] A subscript on effort compensation is dropped to emphasize that the wage payment depends on the agent's effort choice as on the participants judgment decision.

[3] Following the literature on wage dynamic, an infinite horizon is chosen for simplicity reasons. This also reflects that the agent's effort decision exceeds the duration of a contract since she is eager to signal to potential employers in the future.

$$\boldsymbol{e*} := \max_{\boldsymbol{e}} \mathbf{E}[\sum_{t=1}^{\infty} \beta^{1-t} \cdot U_i(W_t, e_{i,t})|y^{t-1}] \tag{6}$$

were $\boldsymbol{W} = (W_1, W_2, ...)$ denotes the principal's strategy function regarding paid compensation and $\boldsymbol{e} = (e_1, e_2, ...)$, the agent's strategy function regarding effort. $\beta < 1$ represents a discounting factor which is assumed to stay constant through time. At the optimal solution ($\boldsymbol{W}^*$, $\boldsymbol{e}^*$), the principal maximizes her lifetime utility with respect to production under the constraint that the agent's lifetime utility exceeds her reservation utility (participation constraint) and the agent maximizes her effort with respect to her decision rule, (Spear and Srivastava, 1987).

2.2 Payment decision under uncertainty

Following Holmström (1999), the agent's wage in period t is based on the expectation of performance in period t conditional on the history of outputs up to that period $y^{t-1} = (y_1, ..., y_{t-1})$,

$$W_t = E[y_t|y^{t-1}] = E[\eta_{i,t}|y^{t-1}] + E[e_{i,t}|y^{t-1}] + E[\epsilon_t|y^{t-1}] \tag{7}$$

Under the assumption of an independently distributed error term with mean 0, $E[\epsilon_t|y^{t-1}] = 0$, and given utility maximizing effort of the agent, wage in period t is solely dependent on $E[\eta_t|y^{t-1}]$, the principal's perception of ability. Harris and Holmström (1982) assume the mean belief about ability to be normally distributed with mean m_t and variance $\sigma_{m,t}$, depending on a prior belief about ability $(m_1, \sigma_{m,1})$. The principal utilizes observations of past output to estimate the value of the agent. The estimate's accuracy increases with the number of periods output is observable. The market's learning process about the agent's ability is subject to the sequence $a_t = \eta_t + \epsilon_t = y_t - e^*_t$, assuming agent i always chooses the optimal effort level. Given the normality and independence assumption on the error term, ϵ_t, and assuming that the dynamic process of ability is a random walk, $\eta_{t+1} = \eta_t + \delta_t$ where $\delta_t \sim i.i.D.$, the market's learning process is well defined.[4] The variance of m_t decreases with time and converges to a steady state in which learning of output observations offsets the

[4]Holmström (1999) defines the learning process as $m_{t+1} = \mu_t * m_t + (1-\mu_t) * a_t$, where $\mu_t = \frac{h_{m,t}}{h_{m,t}+h_\epsilon}$ with h_x being the inverse of the variance of x, a precision term. Moreover, the precision on the learning effect behaves as $h_{m,t+1} = \frac{(h_{m,t}+h_\epsilon)*h_\delta}{h_{m,t}+h_\epsilon+h_\delta}$. See Harris and Holmström (1982) and DeGroot (2005) for a detailed elaboration on the market's learning process.

increased uncertainty of the dynamic development of ability, (Holmström, 1999).

The principal's payment decision given by (7) is the main focus of this thesis. There is a large body of literature, beginning with Tversky and Kahneman (1975), discussing behavioral heuristics in judgment decisions under uncertainty. The moral hazard setting assumes the principal's decision on the appropriate wage to be made under imperfect information about the agent's ability and may therefore be prone to such behavioral effects.

Tversky and Kahneman (1975) identified three key heuristics regarding judgments under uncertainty: representativeness (probabilities are estimated with respect to similarity to stereotypes), availability (probabilities are estimated with respect to information that comes more easily to mind) and adjustment/anchoring (estimates of probabilities are subject to a prior belief). Clearly, these three heuristics are not independent by definition. Information that come more easily to mind may be used to compare probabilities to and may further serve as a prior to adjust from. Hence, Kahneman and Frederick (2002) revised the theory of judgmental heuristics and introduced the concept of *attribute substitution* as the underlying process for cognitive shortcuts affecting judgments under uncertainty. The theory describes the substitution of the target attribute in question by a heuristic attribute if the former is relatively inaccessible. A typical example of attribute substitution is the task of categorical prediction reported in Tversky and Kahneman (1973). Experiment participants were asked to rank the likelihood that a fictive student has specialized in one of nine fields after given a description of the student. The participants reported the same judgments of probability as a control group that ranked the nine fields via similarity to a typical student of the specialization even after discrediting the student's description.[5] Clearly, the participants substituted the target attribute (probability) for an easier available heuristic attribute (similarity).

2.3 Prototype heuristic in wage bargaining

Kahneman and Frederick (2002) discuss prototype heuristic as a generalization of representativeness heuristic and a common method of attribute substitution. The target attribute is (partly) substituted by similarity; the attribute in question is compared to a set of available observations of a set, given the set is homogeneous enough. "The prototype of a set is characterized by the average values of the salient properties of its members," (Kahneman, 2003, p.1463). A variety of experiments provide evidence that prototype heuristic influence willing-

[5] This was done by telling the participants, all graduate students in psychology, that the description had been written while the student was in high school and on the basis of personality tests of dubious validity. The correlation coefficient of the mean judgment of the two groups is 0.98. (Tversky and Kahneman, 1973)

ness to pay decisions (Frederick and Fischhoff, 1998) and categorical prediction (Tversky and Kahneman, 1973) among other fields.[6]

In contract negotiations, principal j is faced with the judgment of agent i's ability under uncertainty, i.e. agent i's ability is unknown and may only be estimated by the observation of the output history. As a consequence, the principal may base her judgment partly on past events that show similarity to the negotiations, may therefore come more easily to mind and serve as a baseline to which the hypothetical outcome may be compared to. In other words, past negotiation outcomes may affect the principals' willingness to pay for agent i's services given that the agent's output history is observed.

The main challenge of this analysis is to establish a measurement of similarity utilized by the average principal in the National Basketball Association to compare athletes and to identify the cluster that may influence a team's judgment on athletes' ability. In general, I will define similarity between two agents as the number of common neighbors they share in a multidimensional space conditional on n characteristic dimensions, where neighbors in R^n are points present in a region of prespecified radius around the point in question. Based on this measure of similarity, one can identify k clusters in R^n so that the similarities in groups are high, while the similarity between groups are low, (Jain, 2010).

The categorization in groups may be utilized by potential employers to compare athletes to each other and to further base their judgment decisions of a specific athlete's ability on observation of similar athletes. Let c be the representative cluster of agent i, a subset of all available observations N, and y_c^{t-1} be the prototype production history for all members of $c \subseteq N$, i.e. the average performance of similar enough athletes.

Assuming prototype heuristic, wage in period t depends on the individual production history and additionally on the prototype history,

$$W_t = E[y_t | y^{t-1}, y_c^{t-1}] \tag{8}$$

If the prototype production history influences wage decisions, $E[y_t|y^{t-1}, y_c^{t-1}] \neq E[y_t|y^{t-1}]$. Hence, the principal's payment decision is influenced by the observation of y_c^{t-1} and may lead to a different wage paid for the agent's production. This effect may either be caused by the principal's perception of agent i's ability, $E[\eta_t|y^{t-1}, y_c^{t-1}]$, or through a difference in observed productivity, given that the agent is a member of cluster c, $E[\epsilon_t|y^{t-1}, y_c^{t-1}] \neq 0$.

[6]See Kahneman and Frederick (2002) for an extensive discussion of empirical evidence for prototype heuristic.

Moreover, this thesis will focus on the average wage of the representative cluster c, that reflects the prototype production history. If NBA teams base their judgment about future performance of an athlete on the prototype of the representative cluster, the cluster's average wage of period $t-1$ should positively affect individual wages after controlling for individual performance.

3 Background: National Basketball Association

The National Basketball Association (NBA) is the men's professional basketball league in North America. 29 out of 30 teams are located in the USA and Toronto hosts the only team outside of the United States. Since 1967, the regular season is 82 games long, currently starts in the last week of October and ends in April of the following year. The league is divided in two conferences, the eastern conference and the western conference whose members are competing for eight playoff spots in each conference. However, each team plays at least two times against every other team in the NBA, independent of conferences. This thesis will concentrate on the regular season performance of athletes. Although the championship is decided in the playoffs, regular season performance is (i) important for playoff performance due to seeding and (ii) important for the individual since it serves as a signal to potential employers.

Since 1970, labour issues in the NBA are governed by a legal contract, the Collective Bargaining Agreement (CBA), between the league and the National Basketball Player's Association (NBPA). The current CBA has been effective since December 8, 2011. The contract determines minimum and maximum salaries for individual athletes, the maximum team payroll (salary cap) and rules regarding player signings, trades, etc. As this paper analyses signed contracts before and after December 8, 2011, I will discuss similarities and differences between the current CBA and the previous one, effective from 2005 to 2011.

In the NBA, athletes that play their first seasons are referred to as *rookies* and athletes whose contracts are expired and are therefore able to sign a new contract are referred to as *free agents*.

3.1 Salary cap

Total team payrolls are restricted by the salary cap defined in the CBA. The National Basketball Association collects total revenues and shares them equally among teams in the hope to ensure competitiveness. Each year's salary cap depends on league-wide projected "Basketball related income" from last year.[7] According to the current CBA, total projected income is multiplied by 44.74% before subtracting projected player benefits and averaged with respect to the number of teams to calculate next year's salary cap. From the 2005/2006 season to the 2010/2011 season, the salary cap was based on 51% of projected Basketball related income. While the salary cap restricts a team's payroll, the CBA contains exceptions which allow teams to sign new players even though the cap is already exceeded or will be exceeded after the signing. These exceptions concern signings of a team's own free agents, replacement of athletes with a season-ending injury/illness and replacement of traded players, (National Basketball Association, 2011, Article VII (6)).

While teams are able to spend more in salaries than the salary cap, teams that exceed a predetermined tax level (higher than the salary cap) are required to pay a tax to the NBA. The tax-rate depends on the incremental team salary above the tax level and if the team exceeded the tax level in three of the four previous seasons. The current tax starts at 150% (250% for "repeater") up to 4,999,999 $ additional spending over tax level and increases in steps of 5 million $. Before the 2013/2014 season, teams paid 100% tax rate on team salary above the tax level, (National Basketball Association, 2011, Article VII (12)(f)).

3.2 Market entry

Market entry in the NBA labour market is governed by a matching process, the NBA draft, which is held prior to the commencement of each NBA season, 10th of July, on a date designated by the league's commissioner. The draft consist of two rounds with the number of selections being equal to the number of teams in the league in each round. The order of selection is determined by the win-loss record of teams of the previous season[8], (National Basketball Association, 2011, Article X (3)) .

No athlete is allowed to sign a contract in the NBA unless he has been eligible for selection in at least one NBA draft: the player has to be at least nineteen years of age during the calendar year the draft is held and at least one NBA season must have elapsed between his high school graduation and the draft in question. If the athlete did not graduate from high

[7] Basketball related Income (BRI) includes items such as broadcast rights, gate receipts, sponsorships, arena naming rights and parking revenues. See National Basketball Association (2011), Article VII(1)(a)

[8] The first draft-round is subject to a lottery process. The fourteen worst teams from the previous season obtain weighted chances to receive a certain selection number. For example, the worst team has a 25% chance to receive the first selection in next year's draft, while the 14th worst team has only a 0.5% chance.

school, at least four calendar years must have elapsed since the graduation of his hypothetical graduation class. International athletes not graduating from an US high school have to be at least twenty-two years old or apply for "Early Entry", by expressing their desire to be selected in the draft in a writing received by the NBA at least sixty days prior to the draft while the player is at least nineteen years old, in order to be eligible for the NBA draft, (National Basketball Association, 2011, Article X (1)).

Once an athlete is selected by a team, a rookie scale contract may be negotiated. All rookie scale contracts with first round selections (currently selections one to thirty) include two guaranteed years with two separate one-year team options for season three and four. The agreed salaries are restricted for the full duration of the rookie contract and decreasing in the pick number. The first pick in 2015 received a first-year salary of 4,753,000 $ while the tenth pick received a first-year salary of 2,068,100 $, (National Basketball Association, 2011, Article VIII (1)). Second round selections (currently selections thirty-one to sixty) do not have a salary scale like first round picks. They are free to negotiate any contract with the team that selected them. All undrafted athletes become unrestricted free agents and are free to negotiate contracts with any team in the NBA.

3.3 Contract background

Once an athlete entered the NBA labour market, he is eligible to sign extensions or new contracts under certain time constraints.[9] Contract length is restricted to a maximum of five years under the current CBA. From the 2005/2006 season to the 2010/2011 season, a contract may have included an additional sixth year for *qualifying veteran free agents*.[10] Depending on the status of the current contract, an athlete may be able to negotiate a contract with any team in the NBA or to negotiate an offer sheet with any team which can be matched by the current employer. The latter situation is referred to as *restricted free agency* and is effective in the fourth year of a rookie scale contract if the team opted to keep the athlete under contract for the third and fourth season or athletes who have been in the league three or fewer years, (National Basketball Association, 2011, Article XI (1)).

Additionally to restrictions on team payroll, athletes' individual salaries are governed by the CBA as well. Both, minimum and maximum salaries are based on the athlete's years of service in the league. In the 2015/2016 season, minimum salaries ranged from 525,093 $

[9] See National Basketball Association (2011), Article VII (7)

[10] The athlete must have played exclusively for one team for the last three seasons. However, if the player has been traded to another team during these three years, he still had the right to sign a contract of six years.

for athletes with zero years of experience up to 1,499,490 $ for athletes with a minimum of ten years of experience in the NBA. See Table 16 in Appendix C for a detailed visualization of minimum salaries from the 2009/10 season to the 2015/16 season, (National Basketball Association, 2011, Article II (6)).

Maximum limits on athletes' salaries are based on the salary cap and previous experience. For a player with up to six years of experience, the greater of 25% of the salary cap or 105% of the player's salary for the last year of the previous contract serves as the upper limit. For athletes between seven and nine years of experience, the greater of 30% of the salary cap or 105% of the player's salary for the last year of the previous contract serves as the upper limit, while a player with a minimum of ten years of experience may be eligible to receive a salary up to 35% of the salary cap or 105% of the player's salary for the last year of the previous contract. See Table 1 for a summary of maximum salaries from the 2009/10 season to the 2015/16 season. Additionally, the CBA 2011 introduced an exception regarding rookie contract extensions. If the athlete meets one of the following "30% Max Criteria" by the time of the contract extension, he is eligible to receive a salary between 25% and 30% of the salary cap: The player has been (i) named to the All-NBA first, second or third team (best fifteen athletes of the regular season voted by journalists) at least two times, (ii) voted an All-Star starter (five most popular athletes according to fan votes) at least two times, or (iii) named NBA MVP (most valuable player voted by journalists) at least once, (National Basketball Association, 2011, Article II (7)) .

Table 1: Maximum annual salary in the NBA

Years in the NBA[11]	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16
0-6	13,520,500$	13,603,750$	12,922,194$	13,668,750$	13,701,250$	14,746,000$	16,407,500$
7-9	16,224,600$	16,324,500$	15,506,632$	16,402,500$	16,441,500$	17,695,200$	19,689,000$
10+	18,928,700$	19,045,250$	18,091,071$	19,136,250$	19,181,750$	20,644,400$	22,970,500$

Notes: Numbers obtained from National Basketball Association (2011) and http://www.cbafaq.com/salarycap.htm on 30th of October, 2016.

3.4 Similarity of NBA athletes

The most obvious and simplest form of categorization in the NBA clusters athletes with respect to their official positions. The positions of the five athletes on the court are point guards (PG), shooting guards (SG), small forwards (SF), power forwards (PF) and centers

[11] An athlete is credited with a year of service as long as he is on team's active or inactive list for at least one day during the season.

(C) and are traditionally responsible for different tasks on the court. The point guard, for example, is the team's ball handler and play maker. He is responsible to lead the team on offense and put teammates in positions to succeed. However, the categorization in positions is very abstract and does not always reflect the contribution of an athlete. Take following example: Player 1 and Player 2 are two individuals in the data set, both primarily playing the point guard position in the 2015/16 season. Selected productivity measurements, displayed in Table 2, show a tremendous difference in the role provided by the two point guards. These performance measurements are 3Par, the share of three-point field goal attempts that are further away from the basket relative to overall field goal attempts; trb%, the percentage of possessions after a missed field goal attempt that the player successfully obtained the ball; ast%, the share of teammates' scoring possession the athlete assisted on; usg%, the percentage of possessions used by the athlete while on the court[12]; pts/g, average points per game scored. While Player 1 uses a majority of team possessions himself (31.6% while he is on the court) and is assisting on almost 50% of teammate's field goals, Player 2 only uses about 15.8% of team possessions but is taking almost 43% of his field goal attempts from three-point area, suggesting that he is more of a recipient of teammate's assists. The third player displayed, Player 3, had quite similar average statistics to Player 2 and one may conclude that based on these characteristics, Player 2 is *more similar* to Player 3 than to Player 1. However, while Player 1 is categorized in the same position as Player 2, point guard, Player 3 plays a completely different position, power forward, whose job description is traditionally very different from that of a point guard. Such heterogeneity among members of positions is the reason I will use different definitions of similarity among athletes to give an alternative to the traditional categorization in terms of positions.

Although it may be that teams compare athletes with similar individuals based on production characteristics, rather than positions, there is already some heterogeneity between position averages. Table 3 displays position averages and standard deviations for the five productivity measurements. The data indicates that power forwards and centers are, on average, superior rebounder with an percentage of available rebounds executed, trb%, of 13.5% and 15.3%, respectively. Wings, shooting guards and small forwards, are traditionally great shooters and scorers which is indicated by their rate of three point attempts with respect to total field goal attempts of 0.373 and 0.356, respectively, and high points per game averages. The average point guard is the team's primary facilitator, indicated by an ast%, the percentage

[12]Possessions end by field goal attempts, approximately 44% of freethrows and turnover. See Section 4.1 for an extensive discussion of the variables used in this thesis.

of team field goals assisted, of 26.5%. While this is a vary narrow picture of the multidimensional performance in NBA games, it still shows that there are heterogeneities between positions and the categorization in positions will be considered as the baseline specification in the econometric analysis and tested against more detailed cluster-specifications.

Table 2: point guard and power forward comparison: An example

	3PAr	**trb%**	**ast%**	**usg%**	**pts/g**
Player 1	0.236	0.124	0.496	0.316	23.475
Player 2	0.425	0.650	0.155	0.158	12.081
Player 3	0.489	0.760	0.114	0.127	7.877

Notes: 3PAr is the fraction of three-point field goals attempts of total field goal attempts ,trb% is the percentage of successfully executed rebounds while on the court, ast% is the percentage of assisted teammate's field goals while on the court, usg% is the percentage of the team's possession ended by the player while on the court and pts/g is the number of scored points per game. Data obtained from http://www.basketball-reference.com on 30$^{\text{th}}$ of October, 2016.

Table 3: Summary statistics - comparison of positions

	3par	**trb%**	**ast%**	**usg%**	**pts/g**	**N**
PG	0.318 (0.150)	0.058 (0.016)	0.265 (0.087)	0.205 (0.048)	9.294 (5.698)	595
SG	0.373 (0.171)	0.065 (0.020)	0.135 (0.069)	0.197 (0.049)	9.107 (5.709)	651
SF	0.356 (0.190)	0.088 (0.025)	0.097 (0.054)	0.177 (0.049)	9.294 (5.698)	625
PF	0.150 (0.194)	0.135 (0.040)	0.083 (0.054)	0.185 (0.048)	8.220 (5.649)	653
C	0.030 (0.087)	0.153 (0.038)	0.072 (0.053)	0.172 (0.055)	7.141 (5.266)	636

Notes: Means calculated over all active players from 2010 to 2016; Standard Deviations in parenthesis. Data obtained from http://www.basketball-reference.com on 30$^{\text{th}}$ of October, 2016.

4 Data and econometric framework

4.1 Data

The data utilized in this article contains performance measurements, player characteristics and salary information for 883 athletes between the 2009/2010 and the 2015/2016 season, yielding an unbalanced panel of 3,159 observations. Performance statistics were collected from Basketball-reference.com and NBA.com. Salary statistics were drawn from ESPN.com, and Basketball-reference.com. Summary statistics are available in Table 4.

I identified two competing individual performance measurements for NBA players, Win Shares (WS) and Value over Replacement Player (VORP), which are publicly available at Basketball-reference.com. Both are based on box score data, a variation of production variables collected by the NBA for every official game, and combine the performance dimensions into one single production variable. The most important difference between the two variables is the method of estimating individual player production with box score variables. While VORP utilizes a +/- method, that is, the positive or negative margin the team performed relative to its opponents while the athlete was contributing, WS is based on techniques introduced by Oliver (2004) to decompose individual production into contribution parts and assign them to the athletes responsible. That being said, VORP and WS are highly correlated in my sample, with a correlation coefficient of 0.912. See Appendix A for a detailed overview of both measurements' calculation.

The box score captures the official statistics collected for every official game by the NBA. A typical game has 48 minutes but some games may exceed this if the score is tied after the regular playing time. Statistics collected and used in this thesis include the amount of games player participated in during the regular season and the average amount of minutes an athlete played per game. Further, the amount of points scored is recorded and the average points per game during a season, along with efficiency estimates (percentage of scoring attempts converted) serve as an approximation of scoring ability. Scoring attempts in the NBA are referred to as field goal attempts and may yield two or three points depending on the distance to the basket if converted successfully. Moreover, the athlete receives two free throws if he is fouled during a field goal attempt or a threshold of fouls is exceeded by the opposing team. Free throws are unopposed attempts close to the basket. The box score also includes assists (a pass by the athlete that leads to a basket), rebounds (gaining possession of the ball after a missed field goal attempt), steals (actively gaining possession of the ball from the opponent), blocks (legally deflecting an opponent's field goal attempt) and turnovers (losing possession of the ball). All the box score variables will be used as season averages in the hope to capture a more robust estimate of an athlete's value.

There is a large body of literature utilizing pure box score data to measure athlete performance such as Kopkin (2012), Stiroh (2007) and Yang and Lin (2012). With the utilization of more advanced productivity measurements, I hope to capture dimensions of performance that are neglected when focusing on simple box score data. WS and VORP are adjusted for the performance of teammates and opponents and account for the pace of games. These adjustments are important since box score variables scale with the amount of possessions and playing time of athletes. Measurements that account for these factors are therefore able to better estimate the individual contribution to aggregated success.

4.2 Econometric framework

The utilization of panel data allows for an analysis of wage effects over time. However, the analysis of individuals over time means that the dependence of observations may affect estimation results and one should therefore control for individual heterogeneity. Since pooled OLS assumes that residuals are independently and identically distributed, standard errors will be incorrect if this dependence is not taken into account. (Moulton, 1986)

Equation (9) describes a simply econometric model based on the theory established in Section 2:

$$W_{i,t} = \kappa + \beta' y_i^{t-1} + \alpha' z_{i,t} + e_{i,t}, \tag{9}$$

where $W_{i,t}$ is individual wage in period t, y_i^{t-1} is the individual production history up to period t, $z_{i,t}$ is a vector of individual time-varying characteristics and $e_{i,t}$ is the estimated error term. Further, I will from now on assume that wages are only affected by one- and two-period lagged production, $y_i^{t-1} = \binom{y_{i,t-1}}{y_{i,t-2}}$.[13] Now, OLS assumes $e_{i,t}$ to be independently and identically distributed. However, the dependence of observations suggests that there are unobserved individual-specific effects affecting wages so that $e_{i,t}$ is not a classical error term and can be further decomposed into

[13]This assumption is due to data constraint. While it is common in wage estimations to only use one-period lagged production, I hope to offer additional information by the inclusion of two lags, since a significant part of contracts in the NBA exceed one year and wages during a contract stay fairly constant.

$$e_{i,t} = \gamma_i + \epsilon_{i,t}, \tag{10}$$

where $\epsilon_{i,t} \sim (0, \sigma_\epsilon^2)$ and γ_i represents such time-invariant individual fixed effects that are not controlled for in a pooled OLS regression. Combining (9) and (10), treating γ_i as $N \times 1$ vector of dummy variables with N being the number of observed individuals, yields the least squares dummy variables (LSDV) estimator for an equation reading as

$$W_{i,t} = \kappa + \beta_1 y_{i,t-1} + \beta_2 y_{i,t-2} + \alpha' z_{i,t} + \lambda' \gamma_i + \epsilon_{i,t} \tag{11}$$

Following Angrist and Pischke (2008), estimating the model with dummy variables is algebraically the same as an estimation in deviations from individual means:

$$W_{i,t} - \bar{W}_i = \beta_1(y_{i,t-1} - \bar{y}_i) + \beta_2(y_{i,t-2} - \bar{y}_i) + \alpha'(z_{i,t} - \bar{z}_i) + (\epsilon_{i,t} - \bar{\epsilon}_i), \tag{12}$$

where the deviations from means eliminates the time-invariant individual fixed effects, γ_i, and the constant, κ. This fixed effects (FE) estimation controls for individual characteristics that may affect estimated coefficients if omitted. The downside of the FE estimation is, however, that possibly interesting information between individuals is filtered out in that it is impossible to estimate effects of time-invariant variables such as race, height, etc.

An alternative to FE estimation is random effects (RE) estimation. RE assumes the individual effects to be randomly distributed in the sample with mean 0, $\gamma_i \sim (0, \sigma_x^2)$, and that γ_i are independent of $\epsilon_{i,t}$. Further, right-side variables, $y_{i,t-1}$, $y_{i,t-2}$ and $z_{i,t}$ are assumed to be independent of both, γ_i and $\epsilon_{i,t}$. These restrictive assumptions are the reason that FE is the method of choice in most panel-data studies. However, if the assumptions hold, RE allows for the estimation of time-invariant effects and should be preferred due to efficiency considerations.

Hausman (1978) proposed a test to choose between RE and FE by comparing the estimated coefficients. Under $H_0 : (e_{i,t} | y_t^{t-1}, z_{i,t}) = 0$, the coefficients of RE and FE should be both consistent and therefore statistically identical. If there is correlation between individual effects and independent variables, between-individual effects may bias the estimation and FE should be the preferred estimation method. The Hausman (1978) test statistic for all specifications displays significant results, which is why I will reject H_0 and use FE as the preferred method for my analysis. Finally, all results reported are estimated with clustered standard errors with respect to individuals to further control for any remaining within individual auto-correlation.

4.2.1 Main variable of interest: prototype wage

I will now develop my estimation strategy to investigate effects of prototype heuristic on wage decisions. The main hypothesis of this thesis is that individual wages are, in addition to past performance, based on past evaluations of performance of similar athletes, represented by their wages. The definition of similarity is key, since categorization is a relative process. Hence, different clustering-specifications will be utilized in the results section.

Let $\{1, ..., i, ..., C\} \in c \subseteq N$ be the representative cluster of individual i, a subset of all individuals N, based on n performance characteristics. Then the average cluster-specific wage,

$$\bar{W}_{c,t} = \frac{\sum_{j=1}^{C} W_{j,t}}{C}, \tag{13}$$

represents the relative market value of cluster c in period t, referred to as prototype wage for individual i in period t. Following the psychological theory established in Section 2, I hypothesize that NBA teams may base their payment decision partly on one-period lagged prototype wage after controlling for individual productivity to estimate the athlete's future production.

The basic econometric wage model considering fixed effects and prototype wage can be written as

$$W_{i,t} = (\kappa + \gamma_i) + \beta_1 y_{i,t-1} + \beta_2 y_{i,t-2} + \delta \bar{W}_{c,t-1} + \alpha z_{i,t} + \epsilon_{i,t} \tag{14}$$

The main hypothesis of this thesis is that $\delta > 0$, i.e. that an athlete's individual wage is positively and significantly affected by one-period lagged average cluster-specific income. Hence, an increase in prototype wage in period $t - 1$ should positively impact individual wages holding individual productivity and time-varying characteristics constant. However, if the wage decision is made exclusively considering individual characteristics, $\delta = 0$ would be expected.

4.3 Cluster analysis

As established in Section 3.4, the simplest form of categorization of NBA athletes is in form of five prespecified positions. However, this specification may neglect a lot of information that would suggest a categorization in more than five groups or based on factors that are not captured by this classification. To account for this, I will utilize an hierarchical clustering algorithm to identify groups of athletes depending on specified productivity characteristics in the hope to offer a robust analysis.

One of the most widely used clustering methods is the k-means algorithm, introduced by MacQueen (1967). The algorithm categorizes N observations in n-dimensional space into K groups so that the similarities of observations within groups are high and the similarities between the group averages are low. Formally, let $X = \{x_{i,j}\}, i = 1, 2, ..., N, j = 1, 2, ..., n$ be the set points in R^n over N observations to be clustered into a set $C = \{c_k\}, k = 1, 2, ...K$ of K clusters. Further, let μ_k be cluster k's centroid. The k-means algorithm minimizes the objective function,

$$J(C) = \sum_{k=1}^{K} \sum_{x_{i,j} \in c_k} \parallel x_{i,j} - \mu_k \parallel^2, \tag{15}$$

the sum of squared residuals between every point in c_k and the cluster specific centroid over all cluster K. The algorithm starts with an initial partition with K clusters, assigns each point to a cluster and calculates the centroids. These steps are repeated until the algorithm converges to a local minimum of the objective function, in that the assignment of new centroids does not change from the previous step. The k-means algorithm requires user input about the number of groups observations are clustered in, K, and the n dimensions clustered on.

In the framework of my research, I will utilize the k-means algorithm to categorize NBA athletes in more or less homogeneous groups with respect to a number of productivity measurements. The goal is to describe the role an NBA player is contributing and to be able to cluster athletes accordingly. Classification is something subjective and vague in nature and it is therefore a difficult task to identify the appropriate specification of dimensions and number of groups used by teams to compare athletes with each other. Hence, the clustering of observations should be tested carefully and rely as less as possible on the researcher's subjective judgment.

In the results section, I will utilize two different clustering specifications with respect to the dimensions categorized on in the hope to offer a robust analysis. According to Jain (2010), the most common method to identify the number of groups categorized in is to compare the within group sum of squared residuals, WSS, of possible number of clusters K given by equation (15). The WSS will naturally decrease as K increases. Hence, I will utilize the proportional reduction of WSS compared to the WSS of the non-clustered data, $WSS(1)$, for every possible cluster solution,

$$\eta_K^2 = 1 - \frac{WSS(K)}{WSS(1)} \tag{16}$$

and the proportional reduction of WSS of each possible cluster solution K compared to the one-step previous solution $K-1$,

$$PRE_K = \frac{WSS(K-1) - WSS(K)}{WSS(K-1)} \tag{17}$$

In the economics literature, the k-means and other hierarchical cluster algorithms have been utilized to identify potentially failing banks, (Alam, Booth, Lee, and Thordarson, 2000), to classify hedge funds based on their investment strategy, (Das, 2003), and to categorize firms according to their coordination strategies in global markets, (Roth, 1992).

5 Results

This section reports results on the econometric model discussed in Section 4. It is critical to first establish a link between an athlete's performance and effort compensation in form of wages. The first subsection will therefore analyze the effect of two alternative performance measurements of NBA athletes, Win Shares (WS) and Value over Replacement Player (VORP), on annual salaries. Subsequent subsections will then test the main hypothesis of this thesis related to prototype wage. I will discuss different specifications of clustering with respect to the dimensions clustered on and the number of groups clustered in. The categorization in positions will serve as a baseline and will be compared to more sophisticated specifications in the subsequent analysis. If NBA teams really do base their payment decision on a comparison of the athlete to his peers, prototype wage may affect individual wages after controlling for individual performance. The econometric model utilized in this thesis is based on Stiroh (2007).

5.1 Individual performance and wages

Following the agency model discussed in Section 2.1, past performance serves as a signal for the athlete's ability and willingness to show effort, thus, positively affecting NBA teams' offered salary. The critical advantage of this analysis is the detailed data on individual performance, in contrast to existing literature on executive pay and productivity that typically investigates the link between individual payment and firm performance.[14] There is an obvious impact of individual performance on joint performance of the team, but I hope to establish a more robust analysis by crediting individuals for their individual production and test incentive effects of NBA contracts.

[14] See Murphy (1999), Coughlan and Schmidt (1985) and Mehran (1995) for example.

Table 4: Variable definitions and descriptive statistics

Variable	Description	Mean(SD)	N
Wage	Annual Salary in 100,000 (2016 $)	55.648 (55.582)	2,854
WS	Estimated Productivity - Win Shares	2.711 (2.894)	3,159
VORP	Estimated Productivity - Value over Replacement Player	0.650 (1.309)	3,159
Age	An athlete's age	26.626 (4.263)	3,159
Experience	The number of season an athlete played in the NBA since 1998	5.879 (4.108)	3,154
Star	Dummy variable: equals 1 if the athlete was names an NBA Allstar, voted by fans and coaches	0.058 (0.233)	3,159
Rookie	Dummy variable: equals 1 if the athlete is still in his rookie contract	0.276 (0.447)	3,159
Games	Number of games played per season	54.596 (23.006)	3,159
Minutes	Number of minutes played per game	20.892 (9.334)	3,159
Points	Number of points per game	8.400 (5.643)	3,159
Rebounds	Number of rebounds per game	3.662 (2.453)	3,159
Assists	Number of assists per game	1.835 (1.804)	3,159
Turnover	Number of turnover per game	1.193 (0.770)	3,159
Steals	Number of steals per game	0.654 (0.426)	3,159
Blocks	Number of blocks per game	0.426 (0.453)	3,159
Ast%	An estimate of teammates' field goals the player assisted for while on the floor	0.128 (0.094)	3,158
Orb%	An estimate of available offensive rebounds successfully executed	0.054 (0.042)	3,158
Drb%	An estimate of available defensive rebounds successfully executed	0.147 (0.062)	3,158
Efg%	An estimate of shooting efficiency accounting for higher value of three-point field goals	0.483 (0.079)	3,151
3PAr	A measure of the player's frequency of three-point field goals	0.244 (0.211)	3,151
Dbpm	Estimated impact on defense	-0.260 (1.942)	3,159
Usg%	An estimate of the percentage of team possessions used by the player	0.187 (0.051)	3,158

Note: Means and standard deviations are calculated by pooling data from the 2009/2010 season to the 2015/2016 season. See VORP calculation for a detailed discussion of estimates.

Table 5 displays fixed effects regression results for the econometric specification,

$$W_{i,t} = (\kappa + \gamma_i) + \beta_1 y_{i,t-1} + \beta_2 y_{i,t-2} + \alpha_1 AGE_{i,t} + \alpha_2 AGE^2_{i,t} + \alpha_3 EXP_{i,t} + \alpha_4 EXP^2_{i,t} + \alpha_5 D_{rookie} + \epsilon_{i,t}, \tag{18}$$

where AGE is the athlete's age in years, EXP is the athlete's experience in years, D_{rookie} is a dummy variable for athletes with rookie contracts and $y_i = WS_i$ in column 1 and $y_i = VORP_i$ in column 2. Both productivity measurements are standardized by subtracting the mean and dividing by their specific standard deviation.[15] Both age and experience positively influence salaries and the effects diminish over time, which is represented by the negative coefficients on age and experience squared. The two variables are obviously highly correlated and age seems to be the dominant factor, while including experience seems to offer additional information regarding individual compensation. Athletes under rookie contracts have a lower base salary due to restrictions explained in Section 3.3 which is reflected by a negative coefficient of about -0.3 on the specific dummy variable. Due to the fact that a significant part of NBA contracts exceeds one season, I consider two lags of individual performance. Both specification 1 and 2 yield positive effects of lagged production on salary. The dependent variable in natural logarithm allows for an interpretation of coefficients as percentage increases ceteris paribus.[16] A one standard deviation increase of WS and VORP increases one-period (two-period) future individual wages by 10.4 % (17.9 %) and 8.8 % (17.6 %), respectively. Column 3 of Table 5 reports estimation results considering both performance measurements. Coefficients on $\text{VORP}_{\text{t-1}}$ and $\text{VORP}_{\text{t-2}}$ show no significant effect on individual salary when controlling for Win Shares, while the coefficients of the remaining variables seem hardly affected.

For model selection, I consider goodness of fit measurements from Akaike (1998),

$$AIC = -2 * ln(L) + k * 2, \tag{19}$$

[15] The productivity measurements are standardized for comparison reasons. Their distribution is quite different with VORP distributed around a mean of 0.650 and WS around a mean of 2.711. The standardization allows for a more or less direct comparison of the coefficients of the two measurements and has no significant effect on the coefficients of the remaining variables.

[16] Technically, this is only valid for coefficients up to 0.10 since $e^x \approx 1 + x$ for $x \in [-0.1, 0.1]$. However, following the literature, this thesis may also interpret coefficients slightly greater than 0.1 in percentage changes of the dependent variable.

and Schwarz (1978),

$$BIC = -2 * ln(L) + k * ln(N), \tag{20}$$

where $ln(L)$ is the maximized log Likelihood of the model, k is the number of parameters estimated and N is the number of observations. Both criteria consider a trade-off between a better fit versus a more parsimonious model. The difference between these two criteria is the "punishment" for a more complex model. Both, AIC and BIC increase with the number of parameters estimated, however, BIC's coefficient on k is larger as long as $N > 7$. As a consequence, the two criteria may disagree when comparing multiple models. In general, a smaller AIC/BIC indicates better fit of the model if the number of observations stays constant across compared models. The model with the smallest criteria should be chosen. AIC and BIC are reported in the last two rows of Table 5. Both criteria favour specification 1 with WS as the productivity measurement, which will therefore be the preferred choice of performance variable in subsequent sections.

5.2 Prototype wage with respect to position

As a first step, I will estimate the effect of one-period lagged prototype wage with respect to positions. Recall, that the NBA categorizes athletes in five positions, point guard (PG), shooting guard (SG), small forward (SF), power forward (PF) and center (C). It is very common that athletes play multiple positions in games and the categorization is according to the position the athlete played the most minutes during the regular season. Figure 3 in Appendix B graphs position-specific average real wages and their development over the sample period. There is rather significant variation in the variable, ranging from \$4.140 million for point guards in the 2009/2010 season up to \$6.783 million for power forwards in the 2012/2013 season. Overall, the mean of the variable is \$5.568 million with a standard deviation of \$664,264 . For the econometric analysis, $\bar{W}_{pos,t-1}$ is defined as the prototype wage in \$1 million.

Column 2 of Table 6 presents fixed-effects results from the following econometric specification:

$$\begin{aligned} W_{i,t} = (\kappa + \gamma_i) + \beta_1 WS_{i,t-1} + \beta_2 WS_{i,t-2} + \delta \bar{W}_{pos,t-1} + \alpha_1 AGE_{i,t} + \alpha_2 AGE^2_{i,t} + \\ \alpha_3 EXP_{i,t} + \alpha_4 EXP^2_{i,t} + \alpha_5 D_{Rookie} + \epsilon_{i,t}, \end{aligned} \tag{21}$$

Table 5: Fixed effects estimation - comparing productivity measurements

	Dependent Variable: Log Salary in 100,000 (2016 $)		
	(1)	(2)	(3)
WS_{t-1}	0.104***		0.110**
	(0.025)		(0.048)
WS_{t-2}	0.179***		0.125**
	(0.030)		(0.054)
$VORP_{t-1}$		0.088***	-0.011
		(0.025)	(0.049)
$VORP_{t-2}$		0.176***	0.063
		(0.031)	(0.055)
AGE	0.645***	0.609***	0.633***
	(0.170)	(0.178)	(0.173)
AGE^2	-0.012***	-0.012***	-0.012***
	(0.003)	(0.003)	(0.003)
EXP	0.117**	0.136**	0.119**
	(0.057)	(0.058)	(0.057)
EXP^2	-0.006*	-0.007**	-0.006*
	(0.004)	(0.004)	(0.004)
D_{rookie}	-0.321***	-0.335***	-0.326***
	(0.087)	(0.088)	(0.087)
Constant	-4.927**	-4.385**	-4.754**
	(2.263)	(2.372)	(2.308)
AIC	1567.627	1578.458	1569.569
BIC	1604.635	1615.466	1617.151

Notes: Standard Errors are clustered with respect to individuals.
Significance levels: 10% : $*$ 5% : $**$ 1% : $***$

where δ is the coefficient of main interest on prototype wage in period $t-1$. Column 2 illustrates that a $1 million increase in prototype wage in period $t-1$ raises an athlete's wage by an average of 6.4% in period t after controlling for individual performance. However, the statistical significance of the coefficient on a 10% level suggests a conservative interpretation of the posititve effect.

Column 3 of Table 6 reports results from the econometric specification (21) extended by year- and position-specific dummies. Although wages are adjusted for salary cap inflation, there could still be some unaccounted time dependency left in the data. Further, I control for any increased base salaries with respect to position since there is evidence for a wage premium for athletes playing the center position in Dey (1997) and a negative fixed effect for forwards in Yang and Lin (2012).

The dummy inclusion slightly increases δ and hardly affects the remaining coefficients. A \$1 million increase in prototype wage in period $t-1$ raises individual wages 7.6% on average, holding individual production constant. The increase of the coefficient decreases its p-value to under 0.05, positively affecting the confidence in the marginal effect of prototype wage. None of the coefficients on year- and position-specific dummies, not reported in Table 6, is statistically significant.

It is important to note that the coefficients on one- and two-period lagged performance are hardly affected at all compared to the specification without prototype wage in column 1. This finding shows, that the positive link between prototype wage and individual income is an additional effect to the positive link between lagged performance and compensation. This suggests that the principal's perception of individual performance is not influenced by the fact that the agent belongs to a certain group of athletes but rather, that the judgment upon future performance is based on the observation of similar athletes additionally to past individual performance. Moreover, the effect is economically relevant for an athlete's salary. As an example, the average wage of power forwards increased from \$5.6 million in the 2011/2012 season to \$6.78 million in the 2012/2013 season. The average power forward negotiating a new contract before the 2012/13 season may receive $0.076 * (6.78 - 5.6) = 8.968\%$ less real salary compared to when an identical athlete would negotiate one year later. Considering that the average power forward earned an annual real salary of \$5,775,505 over the sample period, the prototype wage effect accounts for about \$520,000 less real annual salary if the contract is negotiated prior to the 2012/13 season compared to the 2013/14 season for the average power forward in my sample.

The last two rows of Table 6 report AIC and BIC for the specifications. While AIC suggests that the inclusion of prototype wage with respect to positions positively affects the model's fit, BIC would favor the specification in column 1, without prototype wage. The comparison to the model with year- and position-specific dummy-variables is difficult since the criteria punish a model for the number of parameters estimated and their coefficients show no effect significantly different from zero. That being said, AIC would still favor the specification with dummies in column 3 over the one without dummies in column 2.

Table 6: Fixed effects estimation - prototype wage with respect to position

	Dependent Variable: Log Salary in 100,000 (2016 \$)		
	(1)	(2)	(3)
WS_{t-1}	0.104***	0.103***	0.103***
	(0.025)	(0.025)	(0.025)
WS_{t-2}	0.179***	0.179***	0.179***
	(0.030)	(0.030)	(0.030)
$\bar{W}_{pos,t-1}$		0.064*	0.076**
		(0.035)	(0.039)
AGE	0.645***	0.629***	0.667***
	(0.170)	(0.169)	(0.183)
AGE^2	-0.012***	-0.012***	-0.012***
	(0.003)	(0.003)	(0.003)
EXP	0.117**	0.122**	0.139**
	(0.057)	(0.056)	(0.057)
EXP^2	-0.006*	-0.006*	-0.006*
	(0.004)	(0.004)	(0.004)
$Dummy_{rookie}$	-0.321***	-0.322***	-0.316***
	(0.087)	(0.087)	(0.087)
Constant	-4.927**	-5.150**	-6.334**
	(2.264)	(2.265)	(2.601)
$Dummy_{year}$	No	No	Yes
$Dummy_{position}$	No	No	Yes
AIC	1567.627	1564.287	1559.094
BIC	1604.635	1606.582	1643.684

Notes: Standard Errors are clustered with respect to individuals.
Significance levels: 10% : $*$ 5% : $**$ 1% : $***$

5.3 Prototype wage with respect to classification 1

The previous subsection reported a statistically significant, positive link between individual salary and the lagged average market value with respect to an athlete's positions. However, the coefficient's standard error suggests a conservative interpretation of the effect. For reasons discussed before, the classification in five positions is vague in nature and does not account for the variety within positions. This, as well as the following subsection will try to establish a more robust classification of NBA athletes utilizing the *k*-means algorithm. The algorithm requires input regarding the number of groups clustered in and the dimensions cluster on.

The first specification will utilize simple box score data that is easily accessible and quite visibly describes dimensions of production on the court while having the downside of possibly neglecting some dimensions that may be important for the outcome of basketball games. Six variables, points, assists, rebounds, turnovers, steals and blocks per game were standardized and used to cluster observations in up to fifteen groups. Figure 1 reports the logarithm of within-cluster sum of squares, the proportional reduction of WSS compared to the non-clustered data, η^2, and the proportional reduction of WSS compared to the cluster-solution with one less group clustered in, PRE. There is no clear defined threshold for a "true" or "best" number of groups in the literature. A quite arbitrary technique is the "elbow method" in that the plot of WSS on the number of groups considerably flattens[17], thus, indicating that the marginal contribution of any further increase of K on the reduction of WSS decreases. Due to the lack of a well defined test statistic, this thesis will consider multiple cluster-solutions and I will make sure to outline my reasoning in detail.

Figure 1 shows that neither η^2 nor PRE favour a specific clustering solution. Starting with the categorization of five groups from the official positions as my prior, I will consider the solution $K = 5$ with $PRE = 0.121$. Further, $K = 10$ will be considered as a second cluster-specification, at which we can observe a "bump" in the PRE graph, where the WSS reduces 6% compared to the solution with nine cluster.

Table 7 reports the mean of the six performance variables clustered on for each cluster in the $K = 10$ solution. The categorization in ten groups is able to describe the production of athletes in more detail than the traditional categorization in five groups. For example, both, group 5 and 7 represent "big" positions, power forward and center, indicated by their high rebounding average of 6.056 and 8.463, respectively. However, while group 5 represents defensively minded athletes with high averages in blocked shots per game (1.211), Cluster 7 represents offensively minded player with high points (14.676) and assists per game (1.895) averages. The categorization in positions would not have distinguished between these two types of athletes. Teams, however, might consider it in their hiring decisions.

[17] See Makles (2012).

Table 7: Summary statistics - cluster analysis with respect to points, assists, rebounds, turnovers, steals and blocks per game

	Points	**Assists**	**Rebounds**	**Turnovers**	**Steals**	**Blocks**	**N**
C_1	6.793	1.833	2.145	1.020	0.610	0.149	558
C_2	19.807	3.350	7.139	2.527	1.304	0.785	103
C_3	14.513	1.841	9.568	1.890	0.779	1.968	97
C_4	12.789	4.362	3.005	2.042	0.905	0.186	252
C_5	8.179	1.038	6.056	1.161	0.559	1.211	210
C_6	17.987	7.115	4.256	3.023	1.573	0.328	161
C_7	14.676	1.895	8.463	1.709	0.797	0.709	171
C_8	5.820	0.759	4.013	0.827	0.455	0.535	498
C_9	11.314	2.000	4.050	1.339	1.051	0.368	348
C_{10}	2.779	0.511	1.496	0.447	0.245	0.157	762

Notes: Data obtained from http://www.basketball-reference.com on 30^{th} of October, 2016.

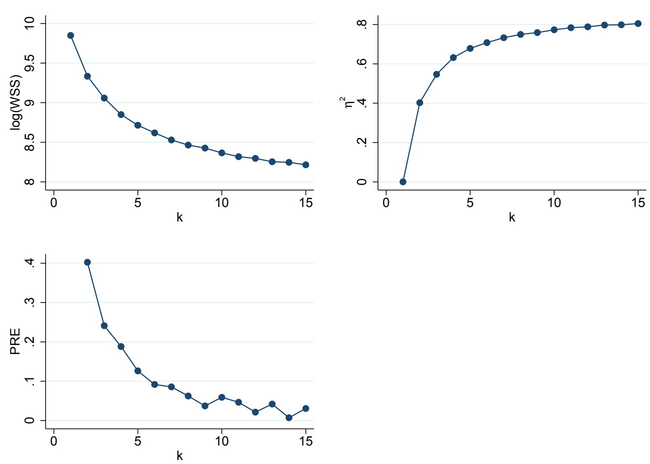

Figure 1: Comparison of possible cluster-solutions dependent on points, assists, rebounds, turnovers, steals and blocks per Game.

Table 8 presents results on the following econometric model:

$$\begin{aligned} W_{i,t} = (\kappa + \gamma_i) + \beta_1 WS_{i,t-1} + \beta_2 WS_{i,t-2} + \delta \bar{W}_{c,t-1} + \alpha_1 AGE_{i,t} + \alpha_2 AGE^2_{i,t} + \\ \alpha_3 EXP_{i,t} + \alpha_4 EXP^2_{i,t} + \alpha_5 D_{Rookie} + \epsilon_{i,t}, \end{aligned} \tag{22}$$

where the prototype wage variable, $\bar{W}_{c,t-1}$, is clustered with respect to points, assists, rebounds, turnovers, steals and blocks per game. Column 1 and 2 present results considering five groups, while column 3 and 4 present results of clustering in ten groups. For a visualization of prototype wage with respect to both specifications and their development over the sample period see Appendix B.

The coefficient on lagged prototype wage is significant and positive for both specifications: $\delta > 0$ in equation (22). Moreover, these findings are robust with respect to the inclusion of cluster- and year-specific dummies. Compared to the results from Table 6 in the previous subsection, the coefficient of one-period lagged productivity is reduced from 0.103 to 0.68 and 0.59 in column 2 and 4, respectively. Hence, a correlation between individual performance and prototype wage is expected in contrast to the specification with respect to positions above. These findings suggest that NBA teams may partly substitute the weight of individual performance in their judgment of future production for the comparison of the athlete to similar agents. However, both effects of one- and two-period lagged productivity remain positive and statistically significant.

The last two rows of Table 8 report the two information criteria, AIC and BIC. Compared to the baseline results in the previous subsection, both cluster-specifications result in smaller AIC. Comparing the two estimations to each other, one should consider the difference of the two criteria in treating increases in the number of parameters estimated. The specification in column 4, where $K = 10$, estimates five additional parameters compared to the specification in column 2, where $K = 5$, as a consequence of the increased number of cluster-dummies. In column 4, only one of the overall nine coefficients on the cluster-dummies is significantly different from 0 and the inclusion hardly affects any other variables. Hence, I would suggest to use AIC as the main criteria when comparing estimations with different number of groups athletes are categorized in since the inclusion of the dummies is due to robustness reasons and seems to hardly affect remaining coefficients estimated.

Table 8: Fixed effects estimation - prototype wage with respect to points, assists, rebounds, turnovers, steals and blocks per game

	Dependent Variable: Log Salary in 100,000 (2016 \$)			
	$K=5$		$K=10$	
	(1)	(2)	(3)	(4)
WS_{t-1}	0.070***	0.068***	0.059**	0.059**
	(0.024)	(0.025)	(0.023)	(0.024)
WS_{t-2}	0.169***	0.171***	0.166***	0.170***
	(0.029)	(0.029)	(0.029)	(0.029)
$\bar{W}_{c,t-1}$	0.037***	0.036***	0.044***	0.042***
	(0.008)	(0.008)	(0.008)	(0.008)
AGE	0.603***	0.554***	0.594***	0.568***
	(0.165)	(0.178)	(0.162)	(0.177)
AGE^2	-0.011***	-0.010***	-0.011***	-0.010***
	(0.003)	(0.003)	(0.003)	(0.003)
EXP	0.121**	0.128**	0.118**	0.125**
	(0.055)	(0.053)	(0.055)	(0.055)
EXP^2	-0.006*	-0.006*	-0.006*	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)
$Dummy_{rookie}$	-0.309***	-0.322***	-0.314***	-0.322***
	(0.086)	(0.085)	(0.085)	(0.085)
Constant	-4.651*	-4.512*	-4.574**	-4.660*
	(2.202)	(2.510)	(2.165)	(2.494)
$Dummy_{year}$	No	Yes	No	Yes
$Dummy_{cluster}$	No	Yes	No	Yes
AIC	1540.543	1521.329	1527.375	1519.753
BIC	1582.838	1605.919	1569.670	1630.778

Notes: Standard Errors are clustered with respect to individuals. Prototype wage, $\bar{W}_{c,t-1}$, is calculated with respect to points, assists, rebounds, turnovers, steals and blocks per game and observations are clustered in five groups in column 1 and 2 and clustered in ten groups in column 3 and 4.
Significance levels: 10% : $*$ 5% : $**$ 1% : $***$

5.4 Prototype wage with respect to classification 2

As mentioned above, simple box score data may miss important aspects of performance that teams look for in their player evaluation. The variables in the previous clustering-specification do not account for any scaling effects in that the average production per game may be purely driven by the amount of minutes played. Hence, I will consider variables that transform box score data in order to account for pace of the game, playing time and other factors that may affect on-court production in this subsection. I will focus on eight dimensions, captured by points per game, assist percentage (ast%), offensive rebounding percentage (orb%), defensive rebounding percentage (drb%)[18], effective field goal percentage (efg%), three-point attempt rate (3PAr)[19], defensive box plus minus (dbpm) and usage percentage (usg%). For the definition of these variables see Table 15 in Appendix C.

Figure 2 reports the logarithm of within-cluster sum of squares, the proportional reduction of WSS compared to the non-clustered data, η^2, and the proportional reduction of WSS compared to the cluster-solution with one less group clustered in, PRE, for the categorization of athletes in up to fifteen groups. Again, there is no evidence of one "best" solution at which the η^2 flattens significantly. However, the PRE graph suggests four possible candidates at which $K \geq 5$ and WSS decreases at least 5% compared to the solution $K-1$: $K=6$ with $PRE = 0.106$, $K = 8$ with $PRE = 0.069$, $K = 11$ with $PRE = 0.075$ and $K = 14$ with $PRE = 0.060$.[20]

Table 9 presents results on the fixed effects estimation in equation (22), considering the four discussed categorization specifications. All four estimations include year- and cluster-specific dummies. The coefficient on prototype wage is statistically significant on a 1% significance level except for the coefficient on prototype wage with respect to fourteen groups, which is significant on a 5% level. Compared to results from Table 8, the coefficient on prototype wage has decreased with the more sophisticated cluster specification. The effect of a $1 million

[18] In contrast to the previous specification, this analysis distinguishes between defensive rebounds, that may be interpreted as an defensive characteristic and offensive rebounds, that secure an additional offensive possession and may therefore be interpreted as an offensive characteristic. Although both, offensive and defensive rebounding, are clearly correlated with body-size, the distinction may offer additional information about an athlete's role. The correlation coefficient of offensive and defensive rebounds per game in my sample is 0.7754.

[19] Three-point attempt rate is a proxy for floor spacing. In modern Basketball, spacing is a very important dimension of team productivity in that it enables easier shots around the basket. The three-point line is at least 6.7 m away from the basket and the less crowded the area around the basket, the easier it is for athletes to get near the basket. In general, field goal percentage is negatively correlated with distance to the basket. Hence, field goals have a higher expected value the less this distance.

[20] In my analyses of solutions up to K=40, K=14 was the last solution to exceed a PRE of 5%.

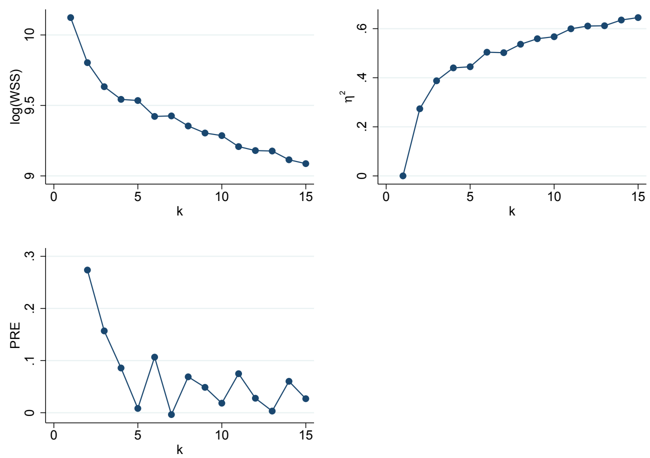

Figure 2: Comparison of possible cluster-solutions dependent on points per game, ast%, orb%, drb%, efg%, 3PAr, dbpm and usg%.

increase in prototype wage in period $t-1$ ranges from 1.8% for $K=14$ in column 4 to as high as 2.6% for $K=11$ in column 3 of Table 8. In contrast to results from subsection 5.3, coefficients on one- and two-period lagged performance are hardly affected by the inclusion of prototype wage. Again, suggesting that information about performance of similar athletes is considered by NBA teams in their judgment decision in addition to information about past individual performance.

The last two rows of Table 9 report the AIC and BIC for all four specifications. For reasons discussed above, I will mainly focus on the AIC results as dummy-inclusion hardly effects the results and estimations differ in the amount of parameters estimated. The information criterion favors the estimation with $K=11$ with an AIC of 1544.398. Compared to the specification exclusively using box score data, the more sophisticated cluster-specification seems to fit the data worse, regardless of the number of groups considered.

Table 9: Fixed effects estimation - prototype wage with respect to points per game, ast%, orb%, drb%, efg%, 3PAr, dbpm and usg%

	Dependent Variable: Log Salary in 100,000 (2016 \$)			
	$K=6$	$K=8$	$K=11$	$K=14$
	(1)	(2)	(3)	(4)
$\mathrm{WS}_{\text{t-1}}$	0.093***	0.092***	0.085***	0.088***
	(0.026)	(0.026)	(0.025)	(0.026)
$\mathrm{WS}_{\text{t-2}}$	0.179***	0.182***	0.172***	0.178***
	(0.030)	(0.030)	(0.030)	(0.030)
$\bar{\mathrm{W}}_{\text{c,t-1}}$	0.025***	0.024***	0.026***	0.018**
	(0.009)	(0.009)	(0.008)	(0.008)
AGE	0.626***	0.623***	0.635***	0.641***
	(0.189)	(0.185)	(0.183)	(0.188)
AGE^2	-0.011***	-0.011***	-0.012***	-0.012***
	(0.003)	(0.003)	(0.003)	(0.003)
EXP	0.127**	0.121**	0.117**	0.112**
	(0.057)	(0.056)	(0.053)	(0.056)
EXP^2	-0.006	-0.005	-0.005	-0.005
	(0.004)	(0.004)	(0.004)	(0.004)
$\mathrm{Dummy}_{\text{rookie}}$	-0.322***	-0.322***	-0.313***	-0.317***
	(0.086)	(0.087)	(0.083)	(0.084)
Constant	-5.216*	-5.294**	-5.267**	-5.195*
	(2.695)	(2.592)	(2.587)	(2.682)
$\mathrm{Dummy}_{\text{year}}$	Yes	Yes	Yes	Yes
$\mathrm{Dummy}_{\text{cluster}}$	Yes	Yes	Yes	Yes
AIC	1554.886	1552.078	1544.398	1549.833
BIC	1639.476	1641.995	1650.136	1666.144

Notes: Standard Errors are clustered with respect to individuals. Prototype wage, $\bar{\mathrm{W}}_{\text{c,t-1}}$, is calculated with respect to points per game, ast%, orb%, drb%, efg%, 3PAr, dbpm and usg% and observations are clustered in 6, 8, 11 and 14 groups in column 1 to 4, respectively.

Significance levels: 10% : $*$ 5% : $**$ 1% : $***$

5.5 Comparing prototype wage specifications

In this section, I will compare the two best fitting specifications from Section 5.2 and 5.3 according to Akaike (1998) and the categorization in positions. Table 10 summarizes AIC results for all previous estimations considering year- and cluster-dummies. The models differ in the specification of the prototype wage variable which is clustered with respect to different dimensions and differs in the number of groups observations are categorized in. Clustering the sample with respect to points, assists, rebounds, turnovers, steals and blocks per game, the categorization in ten groups yielded a lower AIC than the categorization in five groups. Clustering the sample with respect to points per game, ast%, orb%, drb%, efg%, 3PAr, dbpm and usg%, the categorization in eleven groups yielded the lowest AIC. The variable of clustered average wages according to the former specification will be referred to as $\bar{W}_{10,t}$, while the latter will be referred to as $\bar{W}'_{11,t}$. Referring to the average wage with respect to position, I will use the notation $\bar{W}_{pos,t}$.

Table 11 reports results on fixed effects estimations considering different prototype wage specifications. Column 1-3 omit one of the three variables discussed above and column 4 reports results on the estimation including all three variables. $\bar{W}'_{11,t}$ has a significant effect on individual wages only if $\bar{W}_{10,t}$ is omitted in column 2, which may be caused by a positive correlation of the two variables. The coefficient on $\bar{W}_{10,t}$ suggests a significant effect (1% significance level) independent of the other two specifications of prototype wage. Although, coefficients on $\bar{W}_{pos,t}$ are only borderline significant on a 10% level, the positive marginal effect of about 0.065 is robust to the inclusion of other variable-specifications.

Compared to results from previous estimations, the coefficient on $\bar{W}_{10,t}$ remains fairly constant, only slightly decreasing from 0.042 to 0.038 after controlling for the other two prototype wage specifications. The inclusion of $\bar{W}_{10,t}$ seems to negate any significant effects of $\bar{W}'_{11,t}$ on individual wages and the link between prototype wage with respect to position and individual salary stays fairly constant around 0.065. This finding suggests that individual salaries in the NBA are positively affected by information about the athlete's position and the average past compensation of all athletes playing this position even after considering more sophisticated categorizations. However, the categorization of athletes with respect to points, assists, rebounds, turnovers, steals and blocks per game into ten groups yields a highly significant, positive effect on individual wages and, thus, may be the preferred method of NBA teams to compare athletes.

Table 10: Model comparison with respect to AIC

	Positions	Classification 1†		Classification 2‡			
K	5	5	10	6	8	11	14
AIC	1559.094	1521.329	1519.753	1554.886	1552.078	1544.398	1549.833

Notes: Reported AICs are for fixed effects models including year- and cluster-dummies. The models differ in the prototype wage variable-specification. Row one states the classification specification clustered on and row two the number of groups observations are clustered in.
† Observations are clustered with respect to points, assists, rebounds, turnovers, steals and blocks per game.
‡ Observations are clustered with respect to points per game, ast%, orb%, drb%, efg%,3PAr, dbpm and usg%.

Table 11: Fixed effects estimation - prototype wage effect comparison

	Dependent Variable: Log Salary in 100,000 (2016 $)			
	(1)	(2)	(3)	(4)
WS_{t-1}	0.065***	0.092***	0.060**	0.062**
	(0.025)	(0.026)	(0.024)	(0.025)
WS_{t-2}	0.168***	0.171***	0.165***	0.164***
	(0.029)	(0.030)	(0.030)	(0.030)
$\bar{W}_{pos,t-1}$	0.068*	0.065*		0.065*
	(0.036)	(0.037)		(0.036)
$\bar{W}_{10,t-1}$	0.043***		0.038***	0.038***
	(0.008)		(0.008)	(0.008)
$\bar{W}'_{11,t-1}$		0.025***	0.013	0.011
		(0.008)	(0.008)	(0.008)
AGE	0.602***	0.635***	0.617***	0.598***
	(0.177)	(0.181)	(0.176)	(0.176)
AGE^2	-0.011***	-0.012***	-0.011***	-0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
EXP	0.130**	0.129**	0.124**	0.128**
	(0.057)	(0.056)	(0.054)	(0.056)
EXP^2	0.006	-0.006	-0.005	-0.006
	(0.004)	(0.004)	(0.003)	(0.004)
$Dummy_{rookie}$	-0.314***	-0.315***	-0.311***	-0.311***
	(0.085)	(0.086)	(0.084)	(0.084)
Constant	-5,370**	-5.748**	-5.229**	-5.342**
	(2.503)	(2.569)	(2.492)	(2.489)
$Dummy_{year}$	Yes	Yes	Yes	Yes

Notes: Standard Errors are clustered with respect to individuals.
Significance levels: 10% : * 5% : ** 1% : ***

6 Robustness

Section 5 has established a significant effect of prototype wage on individual wages utilizing differing cluster-specifications. Results from fixed effects estimation including multiple prototype wage variables suggest that average wages with respect to points, assists, rebounds, turnover, steals and blocks per game and clustering the sample in ten groups yield the statistically strongest effect. Additionally, past average wages with respect to the five positions seem to offer an additional positive effect on income. This subsection will further analyze the positive effect and test its robustness with respect to alternative performance measurements and lagged individual wages.

6.1 Performance measurement

When interpreting the core results regarding the effects of prototype wage on individual income, a possible source of concern is that the coefficient on prototype wage may be (partly) driven by performance aspects not captured by the chosen production variable. Table 12 presents results on the fixed effects estimation,

$$\begin{aligned} W_{i,t} = \kappa + \beta_1 y_{i,t-1} + \beta_2 y_{i,t-2} + \delta_1 \bar{W}_{pos,t-1} + \delta_2 \bar{W}'_{10,t-1} + \alpha_1 AGE_{i,t} + \\ \alpha_2 AGE_{i,t}^2 + \alpha_3 EXP_{i,t} + \alpha_4 EXP_{i,t}^2 + \alpha_5 D_{Rookie} + \alpha_6 D_{Year} + \gamma_i + \epsilon_{i,t}, \end{aligned} \tag{23}$$

where VORP replaces WS as the productivity measurement, y_i, in column 1 and column 2 reports results on the estimation including two lags of both performance measurements.

The inclusion of VORP has almost no effect on the magnitude of the coefficient on prototype wage nor on the statistical significance of the effect. Moreover, controlling for WS, VORP does not seem to offer any additional information regarding individual wages.

Column 3 of table 12 reports results on the fixed effects estimation with WS as the main performance measurements and controls for All-Star selections of the two-previous seasons. For context, the NBA hosts an All-Star game each season that futures the most popular athletes according to fans and coaches. Past selections may positively influence an athlete's wage even after controlling for individual performance. More popular athletes probably boost game attendance numbers and the overall popularity of the team. Hence, NBA teams may pay a premium for All-Star athletes independent of their on-court production. Moreover, the positive effect of prototype wage may be partly driven by such popularity effects if athletes that belong to certain classifications are simply more likely to be elected to All-star games than others.

Coefficients on both VORP and All-Star selections of the two previous seasons are not significantly different from zero after controlling for WS and prototype wage. Moreover, the effect of prototype wage is robust to the inclusion of the two variables, as presented in column 4 of Table 12.

6.2 Dynamic model

Since prototype wage of cluster c is defined as the average wage of all members of c, the variable may be correlated with lagged individual salaries as long as a significant part of individuals remain in the same classification through multiple seasons. Moreover, since contracts in the NBA often exceed one season, income may stay fairly rigid for a significant subset of the sample and one should control for the auto-correlation of the dependent variable. However, the inclusion of the lagged dependent variable introduces an endogeneity problem and this subsection utilizes a linear generalized method of moments (GMM) to account for any possible bias.

Following Nickell (1981), the inclusion of lagged dependent variables introduces dynamic panel bias additional to the fixed effects problem. To show this, consider the following dynamic fixed effects estimation with no additional parameters,

$$W_{i,t} - \bar{W}_i = \beta(W_{i,t-1} - \bar{W}_i) + (\epsilon_{i,t} - \bar{\epsilon_i}), \tag{24}$$

where $\epsilon \sim i.i.D.$ Demeaning eliminates any individual-specific fixed effects, however, $(W_{i,t-1} - \bar{W}_i)$ is correlated with $(\epsilon_{i,t} - \bar{\epsilon_i})$ since $W_{i,t-1}$ is correlated with $\bar{\epsilon_i}$ by construction and the fixed effects estimation would produce biased coefficients as a consequence, (Baltagi, 2008).

Arellano and Bond (1991) proposed an dynamic model that utilizes an instrumental variable approach in a GMM framework, referred to as "difference GMM" in the literature, to estimate the coefficient of lagged independent variables. Rather than demeaning, one may take first differences to eliminate fixed effects,

$$W_{i,t} - W_{i,t-1} = \beta(W_{i,t-1} - W_{i,t-2}) + (\epsilon_{i,t} - \epsilon_{i,t-1}) \tag{25}$$

Since $W_{i,t-2}$ is highly correlated with $W_{i,t-1} - W_{i,t-2}$, but not with $\epsilon_{i,t} - \epsilon_{i,t-1}$, as long as the error terms are not serially correlated, the authors proposed to use it as an instrument

Table 12: Fixed effets estimation - robustness with respect to productivity

	Dependent Variable: Log Salary in 100,000 (2016 \$)			
	(1)	(2)	(3)	(4)
WS_{t-1}		0.118**	0.067***	0.118**
		(0.048)	(0.025)	(0.048)
WS_{t-2}		0.128**	0.163***	0.123**
		(0.054)	(0.031)	(0.055)
$VORP_{t-1}$	0.040	-0.064		-0.062
	(0.024)	(0.049)		(0.049)
$VORP_{t-2}$	0.161***	0.048		0.048
	(0.030)	(0.054)		(0.054)
$ALL\text{-}STAR_{t-1}$			-0.033	-0.027
			(0.064)	(0.063)
$ALL\text{-}STAR_{t-2}$			0.066	0.063
			(0.078)	(0.078)
$\bar{W}_{pos,t-1}$	0.064*	0.067*	0.068*	0.067*
	(0.036)	(0.036)	(0.036)	(0.036)
$\bar{W}_{10,t-1}$	0.045***	0.044***	0.043***	0.044***
	(0.008)	(0.008)	(0.008)	(0.008)
AGE	0.581***	0.602***	0.595***	0.595***
	(0.183)	(0.179)	(0.177)	(0.179)
AGE^2	-0.011***	-0.011***	-0.011***	-0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
EXP	0.146**	0.130**	0.130**	0.130**
	(0.057)	0.057)	(0.057)	(0.057)
EXP^2	-0.007*	-0.006	-0.006	-0.006
	(0.004)	(0.004)	(0.004)	(0.004)
$Dummy_{rookie}$	-0.328***	-0.319***	-0.317***	-0.322***
	(0.085)	(0.084)	(0.085)	(0.084)
Constant	-5.008*	-5.349**	-5.283**	-5.263**
	(2.584)	(2.534)	(2.507)	(2.541)
$Dummy_{year}$	Yes	Yes	Yes	Yes

Notes: Standard Errors are clustered with respect to individuals.
Significance levels: 10% : $*$ 5% : $**$ 1% : $***$

for $W_{i,t-1} - W_{i,t-2}$. Further, all independent variables with at least two periods in the past, ($W_{i,1}$, $W_{i,2}$,...,$W_{i,t-2}$), may serve as instruments.

Blundell and Bond (1998) proposed a slightly different estimator, referred to as "system GMM estimator" in the literature. This method is an augmented version of the GMM estimator, which utilizes the equation in levels and first differences to obtain additional instruments. This is a useful property for unbalanced panel data-sets since the differentiation of the equation eliminates some observations.

The core principal of GMM is to weight the moments of the two stage least squares (2SLS) estimator to satisfy all moment conditions in a finite sample.[21] The GMM estimator is efficient if the methods are weighted in inverse proportion to their variances and covariances, which, however, are unknown ex-ante. Hence, in a first GMM regression, the moments are weighted by an arbitrary matrix, mostly assuming homosekdasticity. The residuals of the first regression may then be used to construct a matrix that accounts for heteroskedasticity and within-individual covariances in a second estimation, referred to as two-step GMM.

Estimation results on the two-step system GMM are reported in Table 13. Following Blundell and Bond (1998), I will additionally instrument for any possible endogenous control variable, that is, performance and the two prototype wage specifications. The last two rows of Table 13 report the Arellano and Bond (1991) auto-correlation test of an AR(2) process in the error term in first differences[22] and the Hansen (1982) J test of joint validity of the instruments. Instruments are restricted up to a maximum of four lags to avoid over-fitting and in the hope to offer robust Hansen J test-statistics, following Roodman (2009).

Not surprisingly, the inclusion of one-period lagged income affects the coefficient of two-period lagged performance, which is one of the main forces driving the former variable. The coefficient, 0.168 and highly significant in the static model, is not statistically different from zero in the dynamic model. At the same time, the marginal effect of one-period past performance increased dramatically to 0.275 per standard deviation, up from 0.065 in the static model. Regarding the two prototype wage specifications, the average wage with respect to position seems to be driven by an athlete's individual past income. This finding is not too surprising, considering that the position-variable has an auto-correlation coefficient of 0.879. As a consequence, one-period lagged individual wage is very likely to be included in the lagged

[21]For an introduction in linear generalized method of moments estimation see Baum, Schaffer, and Stillman (2003) and for the application in STATA see Roodman (2006).

[22]Tests for AR(1) process in the error term in first differences, not reported here, are significant for all specifications.

prototype wage variable with respect to position and may have been the driving force behind the significant effect in the previous section. The auto-correlation effect of the cluster variable used to calculate $\bar{W}'_{10,t-1}$ of 0.427 is considerably lower.

The significance of the coefficient of $\bar{W}'_{10,t-1}$ is robust to the inclusion of lagged individual income and is increased compared to the static model. A \$1 million increase in prototype wage is followed by an average raise of individual wages of 8.5%. Considering that the average wage in my sample is about \$5.6 million , this effect accounts for an average increase in individual wage of almost \$500,000 per \$1 million in lagged prototype wage.

Column 2 and 3 report results on the restricted estimation, omitting insignificant variables. The coefficient of prototype wage is hardly affected at all. Further, Table 14 in Appendix C reports estimation results on the system GMM model with further instruments, utilizing up to five lags to instrument for lagged income, lagged performance and lagged prototype wage. The findings in this subsection are robust to the inclusion of the additional instruments but the Hansen J statistic decrease as a consequence.[23]

7 Discussion and Conclusions

This thesis is based on an agency model that analyzes contract design when the agent's actions cannot be observed directly by the principal. In this standard model, workers face a prevailing wage and then decide how much effort to supply. The principal's offered compensation depends on estimates of the agent's ability and willingness to shows effort during the contract. Since neither of these variables is directly observable, the principals judges the agent's value based on past output which is imperfectly correlated with effort and ability. The uncertainty of this process allows room for behavioural influences. There is a limited but growing body of literature on psychological consequences for contract design that analyzes wage preferences of agents and principals' judgment bias regarding performance.[24]

This research focuses on the principal's payment decision, assuming that agents always chooses the optimal effort level given an offered compensation.[25] Referring to the theory of

[23]The issue of overfitting in using a large number of instruments and the consequences for the Hansen test of instrument validity are a known problem in the GMM application. See Roodman (2009) for an intensive discussion.

[24]See Camerer and Malmendier (2012) for a literature review on what they call "Behavioral Economics of Organizations".

[25]In reality, agents may have income preferences depending on reference points, shown by Camerer, Babcock, Loewenstein, and Thaler (1997) and Fehr and Goette (2007), and Stiroh (2007) showed that NBA athletes' performance depends on the contract status.

Table 13: Dynamic model - estimation results

	Dependent Variable: Log Salary in 100,000 (2016 $)		
	(1)	(2)	(3)
W_{t-1}	0.475***	0.449***	0.430***
	(0.082)	(0.077)	(0.088)
WS_{t-1}	0.275**	0.265**	0.348***
	(0.117)	(0.116)	(0.107)
WS_{t-2}	0.048	0.049	
	(0.041)	(0.040)	
$\bar{W}_{pos,t-1}$	-0.022	-0.022	
	(0.053)	(0.053)	
$\bar{W}_{10,t-1}$	0.085**	0.092***	0.092***
	(0.033)	(0.032)	(0.034)
AGE	0.233***	0.164**	0.125**
	(0.074)	(0.078)	(0.060)
AGE^2	-0.004***	-0.003**	-0.002**
	(0.001)	(0.001)	(0.001)
EXP	-0.049		
	(0.036)		
EXP^2	0.003		
	(0.002)		
$Dummy_{rookie}$	-0.255***	-0.233***	-0.192***
	(0.065)	(0.064)	(0.050)
$Dummy_{year}$	Yes	Yes	Yes
Instruments	62	60	48
Arellano-Bond AR(2) (p-value)	0.815	0.827	0.313
Hansen J (p-value)	0.319	0.346	0.176

Notes: Standard Errors are clustered with respect to individuals. Estimations are two-stepsystem GMM with the Windmeijer (2005) error correction using the second to fourth lag to instrument for lagged income, lagged performance and lagged prototype wage.
Significance levels: 10% : $*$ 5% : $**$ 1% : $***$

prototype heuristics by Kahneman and Frederick (2002), I argue that NBA teams base their judgment on an athlete's value on a comparison to the evaluated performance of similar athletes in the past. Moreover, athletes are compared to the prototype of their role, represented by average values of salient properties of the homogeneous group an athlete belongs to. This thesis uses different clustering-specifications based on performance measurements to identify athletes that fulfill similar roles on NBA teams. The average real wage of an athlete's group serves as a measurement of relative market value of the cluster since salaries are deflated with respect to the league's salary cap, the maximum amount teams are allowed to spend on player salaries, thus, controlling for any effects attributed to wage inflation. The main hypothesis of this thesis is that an athlete's individual wage is positively driven by past prototype wage of the agent's representative cluster after controlling for past individual performance and individual characteristics such as age and experience.

Utilizing fixed effects models, results show a statistically significant effect of one-period lagged prototype wage after controlling for past individual performance. Moreover, the comparison of different cluster-specification yielded two types of categorizations NBA teams base their comparison of athletes on. First, NBA teams seem to compare athletes according to their official positions; a one million US Dollar increase in one-period lagged position-specific average wage raises individual wages 6.8% after controlling for the positive effect of individual performance. However the variable's coefficient is only borderline significant on a 10% level. The second identified classification of athletes is according to points, assists, rebounds, turnovers, steals and blocks per game into 10 distinct groups. Individual wages increase 4.3% per $1 million in one-period lagged cluster-specific average wage. The coefficient is highly significant on a 1% level. These findings offer evidence that NBA teams' payment decisions are positively linked to past season's average wage of an athlete's role and, thus, indirectly on past average performance of the role independent of the fact how much value the athlete may offer the team.

The findings are robust to the inclusion of other performance measurements, other prototype wage specifications and possible popularity effects measured in past All-Star selections. However, the positive effect of prototype wages with respect to positions is not robust to the inclusion of one-period lagged individual wage in a system GMM model. The coefficient turns negative but is not statistically different from zero. Thus, indicating that the prototype wage effect with respect to positions is mostly driven by the lagged dependent variable. The positive effect of one-period lagged prototype wage with respect to points, assists, rebounds, turnovers, steals and blocks per game on individual wages is robust to the inclusion of lagged individual wages and, further, increases to 9.2% per \$1 million. These findings suggests that (i) NBA teams may neglect the simplistic classification in positions and categorize athletes in more than five groups considering heterogeneities within positions and (ii) the conditional effect of past market value on agreed salaries in wage bargaining may be even stronger than expected by the fixed-effects results, assuming that individual wages are correlated over time even after controlling for individual-fixed effects.

The main limitation of this thesis is the utilized clustering method. The performance-dimensions used to categorize observations have to be chosen by the researcher and, thus offering additional room for biases. Moreover, the literature does not offer satisfying quantitative criteria to identify the "best" number of groups observations are clustered in.[26] With the consideration of multiple specifications of the k-means algorithm, I hope to offer robust results but cannot rule out that NBA teams classify NBA athletes based on other performance dimensions, or that they compare athletes with each other in more detail, resulting in an increased number of groups athletes are clustered in. Further, teams may not compare NBA athletes to the prototype of his role but rather to the most similar athlete. Therefore, future research may consider the degree of similarity between athletes based on their euclidean distance to test this alternative hypothesis.

Another point of concern is that annual salaries during a contract are auto-correlated to a certain extend, independent of the athlete's performance during the contract. Unfortunately, contract status is not available in my database. Future research may analyze contract value at the time of signing as the dependent variable instead of annual salary to test the robustness of this thesis' findings. The dynamic model considers one- and two-period lagged salary and, hence, partly controls for the auto-correlation.

This research analyzes the link between relative market value of categorizations of agents in a setting that offers publicly accessible performance measurements of rather high quality. One can easily imagine settings where individual performance is harder to measure, where employers have less ability to decompose the causal effect of performance and the design of incentive contracts is therefore even more challenging. In such settings, the consequences of moral hazard are likely to be even worse in that the uncertainty about an employee's value for the employer increases. Hence, I would argue that this research provides rather robust evidence of a predictable inefficiency in contract designs under uncertainty, which may even increase with difficulty for the principal to identify the causal effects on individual

[26]For the interested reader see (Jain, 2010) on a detailed overview of clustering methods in data analysis and inherent unsolved problems regarding these methods.

performance.

References

Akaike, H. Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer, 1998.

Alam, P., Booth, D., Lee, K., and Thordarson, T. The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks: an experimental study. *Expert Systems with Applications*, 18, pages 185–199, 2000.

Angrist, J. D. and Pischke, J.-S. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

Arellano, M. and Bond, S. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58, pages 277–297, 1991.

Baltagi, B. *Econometric analysis of panel data*. John Wiley & Sons, 2008.

Barber, B. M. and Odean, T. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21, pages 785–818, 2008.

Baum, C. F., Schaffer, M. E., and Stillman, S. Instrumental variables and gmm: Estimation and testing. *Stata Journal*, 3, pages 1–31, 2003.

Bertrand, M. and Mullainathan, S. Are ceos rewarded for luck? the ones without principals are. *Quarterly Journal of Economics*, pages 901–932, 2001.

Blundell, R. and Bond, S. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87, pages 115–143, 1998.

Camerer, C. and Malmendier, U. Behavioral economics of organizations. *Behavioral Economics and Its Applications*, page 235, 2012.

Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. Labor supply of new york city cabdrivers: One day at a time. *Quarterly Journal of Economics*, pages 407–441, 1997.

Coughlan, A. T. and Schmidt, R. M. Executive compensation, management turnover, and firm performance: An empirical investigation. *Journal of Accounting and Economics*, 7, pages 43–66, 1985.

Das, N. Hedge fund classification using k-means clustering method. In *9th International Conference on Computing in Economics and Finance*, pages 11–13, 2003.

DeGroot, M. H. *Optimal statistical decisions*, volume 82. John Wiley & Sons, 2005.

Dey, M. S. Racial differences in national basketball association players' salaries: A new look. *The American Economist*, 41, pages 84–90, 1997.

Fehr, E. and Goette, L. Do workers work more if wages are high? evidence from a randomized field experiment. *American Economic Review*, 97, pages 298–317, 2007.

Frederick, S. and Fischhoff, B. Scope (in) sensitivity in elicited valuations. *Risk Decision and Policy*, 3, pages 109–123, 1998.

Grossman, S. J. and Hart, O. D. An analysis of the principal-agent problem. *Econometrica*, pages 7–45, 1983.

Hansen, L. P. Large sample properties of generalized method of moments estimators. *Econometrica*, pages 1029–1054, 1982.

Harris, M. and Holmström, B. A theory of wage dynamics. *Review of Economic Studies*, 49, pages 315–333, 1982.

Hart, O. D. and Holmström, B. *The theory of contracts.* Department of Economics, Massachusetts Institute of Technology, 1986.

Hausman, J. A. Specification tests in econometrics. *Econometrica*, pages 1251–1271, 1978.

Holmström, B. Moral hazard and observability. *Bell Journal of Economics*, pages 74–91, 1979.

Holmström, B. Managerial incentive problems: A dynamic perspective. *Review of Economic Studies*, 66, pages 169–182, 1999.

Jain, A. K. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31, pages 651–666, 2010.

Kahneman, D. A perspective on judgment and choice: mapping bounded rationality. *American Psychologist*, 58, pages 697, 2003.

Kahneman, D. and Frederick, S. Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and Biases: The Psychology of Intuitive Judgment*, 49, 2002.

Kliger, D. and Kudryavtsev, A. The availability heuristic and investors' reaction to company-specific events. *Journal of Behavioral Finance*, 11, pages 50–65, 2010.

Kopkin, N. Tax avoidance how income tax rates affect the labor migration decisions of nba free agents. *Journal of Sports Economics*, 13, pages 571–602, 2012.

Lee, B., O'Brien, J., and Sivaramakrishnan, K. An analysis of financial analysts' optimism in long-term growth forecasts. *Journal of Behavioral Finance*, 9, pages 171–184, 2008.

MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, pages 281–297. Oakland, CA, USA., 1967.

Makles, A. Stata tip 110: How to get the optimal k-means cluster solution. *Stata Journal*, 12, pages 347, 2012.

Mehran, H. Executive compensation structure, ownership, and firm performance. *Journal of Financial Economics*, 38, pages 163–184, 1995.

Moulton, B. R. Random group effects and the precision of regression estimates. *Journal of Econometrics*, 32, pages 385–397, 1986.

Murphy, K. J. Executive compensation. *Handbook of labor economics*, 3, pages 2485–2563, 1999.

National Basketball Association. *NBA collective bargaining agreement*. 2011.

Nickell, S. Biases in dynamic models with fixed effects. *Econometrica*, pages 1417–1426, 1981.

Oliver, D. *Basketball on paper: rules and tools for performance analysis*. Potomac Books, Inc., 2004.

Roodman, D. How to do xtabond2: An introduction to difference and system gmm in stata. *Center for Global Development working paper*, 2006.

Roodman, D. A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71, pages 135–158, 2009.

Roth, K. International configuration and coordination archetypes for medium-sized firms in global industries. *Journal of International Business Studies*, 23, pages 533–549, 1992.

Schwarz, G. Estimating the dimension of a model. *Annals of Statistics*, 6, pages 461–464, 1978.

Spear, S. E. and Srivastava, S. On repeated moral hazard with discounting. *Review of Economic Studies*, 54, pages 599–617, 1987.

Stiroh, K. J. Playing for keeps: Pay and performance in the nba. *Economic Inquiry*, 45, pages 145–161, 2007.

Tversky, A. and Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, pages 207–232, 1973.

Tversky, A. and Kahneman, D. Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making*, pages 141–162. Springer, 1975.

Windmeijer, F. A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of Econometrics*, 126, pages 25–51, 2005.

Yang, C.-H. and Lin, H.-Y. Is there salary discrimination by nationality in the nba? foreign talent or foreign market. *Journal of Sports Economics*, 13, pages 53–75, 2012.

8 Appendices

A Data

This section will give a brief overview of both, the WS and VORP calculations, based on the articles available at Basketball-reference.com. Both calculations utilize individual-specific and team-specific variables. I will refer to the former with the subscript i and to the latter with the subscript tm. Further, league-wide averages are referred to with the subscript lg.

Value over Replacement Player (VORP)

Developed by Daniel Myers[27], VORP measures athlete productivity per share of minutes played during a season. The basis for the calculation is a statistical Plus-Minus Method, Box Plus Minus (BPM), which estimates the marginal productivity of a player per 100 possessions in comparison to league average. For example, a Team with a mean BPM of +4 outscores an average NBA Team 4 points per 100 possessions.

The basis for the BPM calculation is an regularized adjusted Plus-Minus (RAPM)[28] sample from the 2000/2001 to the 2013/2014 season. A variety of box-score based variables are chosen as independent variables in a regression to estimate there impact on RAPM and following coefficients were estimated:

$$\begin{aligned} BPM_i = 0.123 * MPG_i + 0.120 * ORB\%_i - 0.151 * DRB\%_i + 1.256 * STL\%_i + \\ 0.532 * BLK\%_i - 0.306 * AST\%_i + 0.921 * (TOV\%_i * USG\%_i) + \\ 0.711 * Scoring_i + 0.726 * (AST\%_i * TRB\%_i)^{1/2}, \end{aligned} \tag{26}$$

where $Scoring_i = USG\%_i * (1 - TOV\%_i) * [2 * (TS\%_i - TS\%_{tm}) + 0.017 * AST\%_i + 0.298 * (3PAr_i - 3PAr_{lg}) - 0.213]$.[29]

[27]See www.basketball-reference.com/about/bpm.html

[28]RAPM is a lineup based variable of individual production. Every possession of a NBA Game futures ten observations, five players of team A and five players of team B. In a sequence of possessions, not interrupted by a substitution, these ten independent variables affect the dependent variable Margin = points scored of team A - points scored of team B. An athletes individual contribution is then calculated by solving the system of equations over one/multiple seasons. Additionally, ridge-regression (Hoerl 1962) is utilised to account for outliers as a consequence vast differences in minutes on the field. For the interested reader, see http://www.82games.com/comm30.htm for the original article on adjusted Plus-Minus.

[29]See basketball-reference.com/about/bpm.html for an extensive Discussion about the independent variables

BPM data is adjusted by playing time to receive an estimate of value generated by an athlete. Since BPM is centered around league average at 0, athletes below average would produce a negative value. Hence, the variable is adjusted to represent value produced relative to a theoretical "replacement level player" defined as an athlete on minimum salary and/or an athlete not part of a team's regular rotation, whose BPM is estimated to be -2.0[30] An individual's contribution during a season is calculated as $VORP_i = [BPM_i - (-2.0)] * T_i$, where T_i represents the individual share of total available playing time.

Win Shares (WS)

Win Shares are based on methods developed by Oliver (2004) and is defined as the sum of offensive and defensive WS to account for value added while the own team is in possession of the ball and while the opponent is.

Offensive WS are based on

$$\begin{aligned} PointsProduced_i = (FGPart_i + ASTPart_i + FTPart_i) * (1 - \frac{OR_{tm}}{ScPoss_i} * \\ ORweight_{tm} * Play\%_i) + ORPart_i, \end{aligned} \tag{27}$$

where FG Part, AST Part, FT Part and OR Part are partial credits for field goals scored, field goals assisted, free throws converted and offensive rebounds, respectively[31], TmPlay% is the percentage of possessions the Team scored and $ORweight_{tm}$ is an estimate of the value of offensive rebounds.[32] Individual Points Produced are then compared to expected Point Produced, given by $0.92 * PPP_{lg} * Poss_i$, where PPP_{lg} are the league-wide average points per possession and $Poss_i$ is the number of individual offensive possessions.[33] The difference yields the athlete's marginal offense which is divided by marginal points per win, $PPG_{lg} * \frac{Pacetm}{Pacelg} * \frac{1}{3}$,

chosen for the regression. And table 15 for a description of the variables with the exception of MPG, which is the average amount of minutes per game played by the athlete.

[30]The definition of replacement level is especially dificult in the NBA since there is (limited) incentive to hire athlete's below replacement levels due to development reasons in the case of very young players or tactical reasons due to the setup of the draft process of the NBA.

[31]Field goals, for example, are often assisted by other athletes. The calculations are based on estimates of the value of assists, which is subtracted from the scorer's credits and reflected in the assisting athlete's AST Part. For the detailed calculations see Oliver (2004) Appendix 1.

[32]$ORweight_{tm} = \frac{(1-OR\%_{tm})*Play\%_{tm}}{(1-OR\%_{tm})*Play\%_{tm}+OR\%_{tm}*(1-Play\%_{tm})}$

[33]Expected Points Produced are weighted by 0.92 to adjust it to "replacement level", hence, to assure that no significant sub sample of total athletes has negative Points Produced.

where PPG_{lg} are the league's average points per game, $Pace_{tm}$ is the number of possessions per 48 minutes and $Pace_{lg}$ the league's average number of possessions per 48 minutes. The multiplication of 1/3 is due to the definition of three Win Shares being equal to one win. An athlete's offensive WS are then calculates as "marginal offense" divided by "marginal points per win".

The core of Defensive WS is an athlete's individual defensive Rating (DRtg) developed by Oliver (2004),

$$DRtg_i = DRtg_{tm} + 0.2 * [100 * OPPP_{tm} * (1 - Stop\%_i) \\ -DRtg_{tm}], \tag{28}$$

where $OPPP_{tm}$ is the number of points scored by the opponent per scoring possession, $DRtg_{tm}$ is the team specific Defensive Rating, defined as points allowed per 100 possessions and $Stop\%$ is an estimate of the rate an athlete forces a defensive stop.[34] The formula accounts for the fact that Defense is a team effort in Basketball and adjusts the team's defensive rating according to individual success while assuming that defensive Possessions are smoothly distributed among the five player on the court (0.2 coefficient). Marginal Defensive (MDef) is then calculated by comparing individual defensive rating to expected one given by the league average:

$$MDef_i = \frac{MP_i}{MP_{tm}} * DPoss_{tm} * (PPP_{lg} * 1.08 - \frac{DRtg_i}{100}), \tag{29}$$

where the coefficient of 1.08 is simply an adjustment for "replacement level" to assure that no significant sub population has negative marginal Defense. As for offensive WS, the marginal defense is divided by marginal points per win, $PPG_{lg} * \frac{Pace_{tm}}{Pace_{lg}} * \frac{1}{3}$, to calculate defensive WS.

Total individual WS is simply the sum of offensive and defensive WS. The variable is designed as to estimate individual contribution to team wins. Hence, a individual WS of athletes should add up approximately to the team's regular season wins. Accoridng to basketball-reference.com[35], the root mean squared error of this comparison is 3.41 wins since the 1962-63 season.

[34] $Stop\%_i = \frac{Stops_i * MP_j}{DPoss_{tm} * MP_i}$ where MP_j is the number of minutes played by the opposing player j, $DPoss_{tm}$ is the number of team-specific possession on defense and Stops are a function of steals, blocks and defensive rebounds. See Oliver (2004) for an extensive explanation.

[35] See http://www.basketball-reference.com/about/ws.html

B Figures

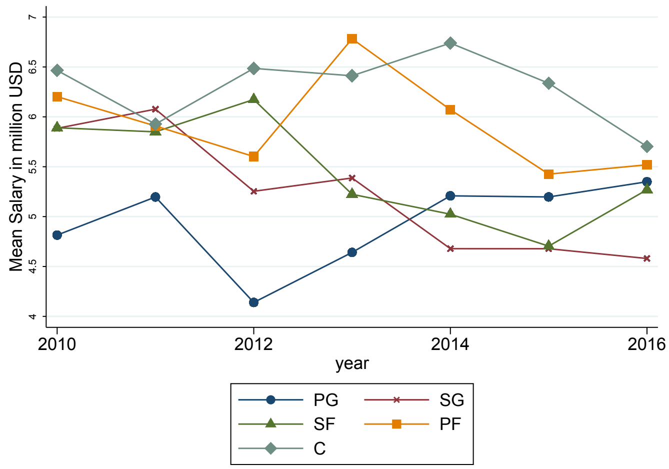

Figure 3: Average salaries with respect to position

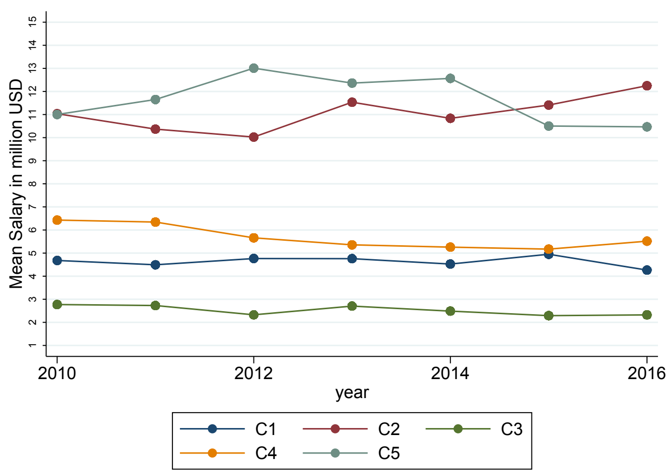

Figure 4: Average salaries with respect to points, assists, rebounds, turnover, steals and blocks per game; $K = 5$

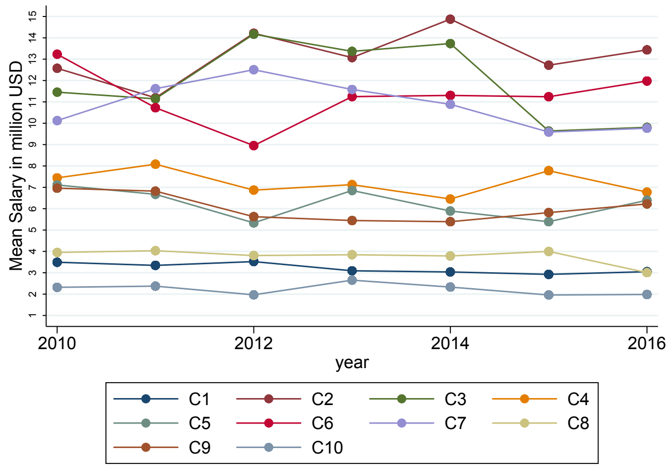

Figure 5: Average salaries with respect to points, assists, rebounds, turnover, steals and blocks per game; $K = 10$

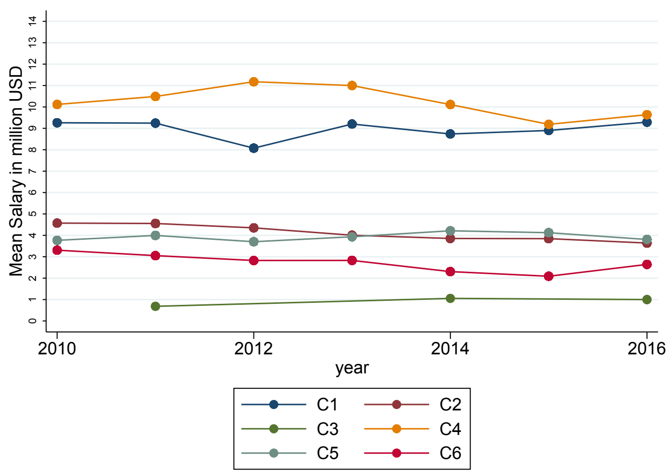

Figure 6: Average salaries with respect to points per game, ast%, orb%, drb%, efg%, 3PAr, dbpm and usg%; $K = 6$

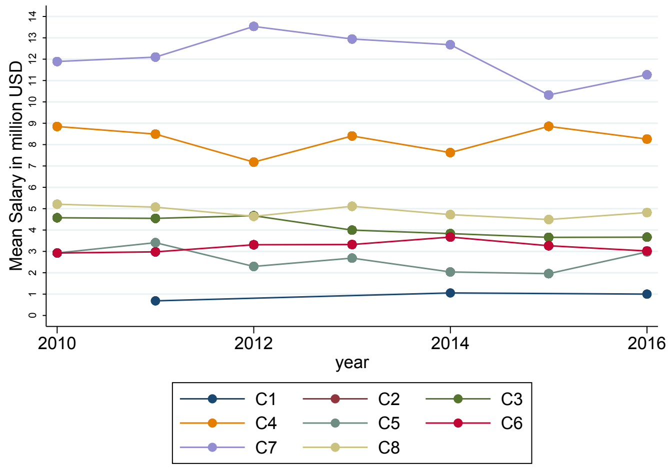

Figure 7: Average salaries with respect to points per game, ast%, orb%, drb%, efg%, 3PAr, dbpm and usg%; $K = 8$

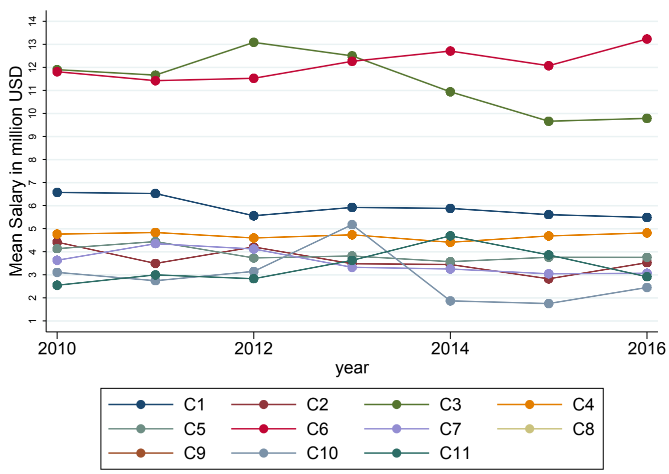

Figure 8: Average salaries with respect to points per game, ast%, orb%, drb%, efg%, 3PAr, dbpm and usg%; $K = 11$

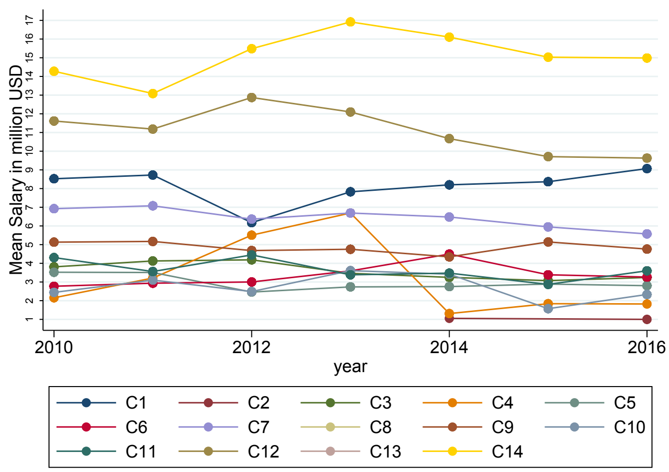

Figure 9: Average salaries with respect to points per game, ast%, orb%, drb%, efg%, 3PAr, dbpm and usg%; $K = 14$

C Tables

Table 14: Dynamic model - Robustness with respect to the number of instruments

	Dependent Variable: Log Salary in 100,000 (2016 \$)		
	(1)	(2)	(3)
$W_{t\text{-}1}$	0.506***	0.470***	0.421***
	(0.078)	(0.074)	(0.086)
$WS_{t\text{-}1}$	0.277**	0.273**	0.364***
	(0.116)	(0.112)	(0.106)
$WS_{t\text{-}2}$	0.059	0.059	
	(0.042)	(0.041)	
$\bar{W}_{pos,t\text{-}1}$	-0.010	-0.010	
	(0.052)	(0.052)	
$\bar{W}_{10,t\text{-}1}$	0.079**	0.086***	0.092***
	(0.033)	(0.032)	(0.034)
AGE	0.213***	0.145*	0.122**
	(0.073)	(0.076)	(0.060)
AGE^2	-0.004***	-0.003**	-0.002**
	(0.001)	(0.001)	(0.001)
EXP	-0.055		
	(0.037)		
EXP^2	0.003		
	(0.002)		
$Dummy_{rookie}$	-0.240***	-0.221***	-0.190***
	(0.065)	(0.063)	(0.049)
$Dummy_{year}$	Yes	Yes	Yes
Instruments	66	64	51
Arellano-Bond AR(2) (p-value)	0.736	0.768	0.335
Hansen J (p-value)	0.192	0.225	0.132

Notes: Estimations are two-step system GMM with the Windmeijer (2005) error correction using the second to fifth lag to instrument for lagged income, lagged performance and lagged prototype wage.
Significance levels: 10% : $*$ 5% : $**$ 1% : $***$

Table 15: Transformed variables

Variable	**Definition**	**Description**
PT_i	$\frac{MP_i}{MP_{tm}/5}$	Estimated share of total Playing Time.
$TRB\%_i$	$100 * \frac{TRB_i}{PT_i*(TRB_{tm}+TRB_{opp})}$	An estimate of available total rebounds successfully executed.
$ORB\%_i$	$100 * \frac{ORB_i}{PT_i*(ORB_{tm}+DRB_{opp})}$	An estimate of available offensive rebounds successfully executed.
$DRB\%_i$	$100 * \frac{DRB_i}{PT_i*(DRB_{tm}+ORB_{opp})}$	An estimate of available defensive rebounds successfully executed.
$STL\%_i$	$100 * \frac{STL_i}{PT*Poss_{opp}}$	An estimate of opponents' possessions that ended in a steal by the player.
$BLK\%_i$	$100 * \frac{BLK_i}{PT*(FGA_{opp}-3PA_{opp})}$	An estimate of opponents' two-point field goal attempts blocked by the player.
$AST\%_i$	$100 * (\frac{AST_i}{PT_i*FG_{tm}} - FG_i)$	An estimate of teammates' field goals the player assisted for.
$USG\%_i$	$100 * \frac{FGA_i+0.44*FTA_i+TOV_i}{PT_i*(FGA_{tm}+0.44*FTA_{tm}+TOV_{tm})}$	An estimate of the percentage of team possessions used by the player.
$TOV\%_i$	$100 * \frac{TOV_i}{FGA_i+0.44*FTA_i+TOV_i}$	An estimate of Turnovers per 100 possessions.
$TS\%_i$	$\frac{PTS_i}{2*(FGA_i+0.44*FTA_i)}$	A measure of shooting efficiency including three-point field goals and free throws.
$3PAR_i$	$\frac{3PA_i}{FGA_i}$	A measure of the player's frequency of three-point field goals.

Table 16: Minimum anual salary in the NBA

Years in the NBA	**2009/10**	**2010/11**	**2011/12**	**2012/13**	**2013/14**	**2014/15**	**2015/16**
0	457,588$	473,604$	473,604$	473,604$	490,180$	507,336$	525,093$
1	736,420$	762,195$	762,195$	762,195$	788,872$	816,482$	845,059$
2	825,497$	854,389$	854,389$	854,389$	884,293$	915,243$	947,276$
3	855,189$	885,120$	885,120$	885,120$	916,099$	948,163$	981,348$
4	884,881$	915,852$	915,825$	915,852$	947,907$	981,084$	1,015,421$
5	959,111$	992,680$	992,680$	992,680$	1,027,424$	1,063,384$	1,100,602$
6	1,033,342$	1,069,509$	1,069,509$	1,069,509$	1,106,942$	1,145,685$	1,185,784$
7	1,107,572$	1,146,337$	1,146,337$	1,146,337$	1,186,459$	1,227,985$	1,270,964$
8	1,181,803$	1,223,166$	1,223,166$	1,223,166$	1,265,977$	1,310,286$	1,356,146$
9	1,187,686$	1,229,255$	1,229,255$	1,229,255$	1,272,279$	1,316,809$	1,362,897$
10+	1,306,455$	1,352,181$	1,352,181$	1,352,181$	1,399,507$	1,448,490$	1,499,187$