NBA Contracts and Recency Bias:
An Investigation into Irrationality in Performance Pay Markets

Casey Fox

Department of Economics

Haverford College

Advisor: Richard Ball

Spring 2015

Abstract:

This paper examines the impact of lagged performance on free agent contracts for players in the National Basketball Association. The main approach of the paper is twofold. The first piece investigates how past performance affects future performance in the two seasons after contract year and compares it to the impact previous performance has on contract terms for free agent players. The second piece investigates the rationality of free agent contracts in their entirety by comparing the impact of lagged performance on total accumulated production and total dollar value paid. The goal is to determine if performance prior to contract year is underweighted in contract decision-making relative to its predictive power of future performance. There is evidence that performance in years prior to contract year is overlooked in contract determination decisions by NBA general managers, and there is mild evidence that performance data two years prior to contract year are underweighted given their predictive power of future performance.

Table of Contents

Introduction

People’s behavior does not conform to strict rationality as traditional economics often assumes. Recent research in psychology and economics indicates that people often rely on rules of thumb or heuristics to evaluate the likelihood of events when economics expects them to make calculated decisions based upon all available data and evidence. A heuristic is a technique which humans use to find the solution to a complicated question using a simpler, experience based approach to estimate the solution, thereby easing the cognitive load on the brain. Kahneman (2011) describes a number of cognitive biases such as recency bias and the availability heuristic that affect the process of rational decision making involving statistical reasoning. Further, there is evidence that these biases occur even when the costs of incorrect assessment are high.

In markets where pay is based on performance, cognitive biases can lead to significant market inefficiency as many employees will be paid in a manner inconsistent with their true ability and performance. Firms will incorrectly value their employees causing payroll misallocation. Recency bias manifests itself when a manager attempts to evaluate an employee’s performance and overlooks their past performance at the expense of their very recent performance in estimating the employee’s skill or talent level while determining their future salary or bonus. The availability heuristic can also affect this potential for inefficiency as it states that decision makers are expected to overweight those events which are more memorable, which can be more recent or more extraordinary, in approximating the likelihood with which an event occurred. For example, a decision maker may remember an impressive period of performance for an individual and overlook their overall performance which may have been well below average outside that impressive period.

Each year NBA teams spend millions of dollars on free agents in an attempt to secure the best talent for their own team. The assumption made in this paper is that each team aspires to win games with the goal of winning the NBA championship as economic research done by Andrew Healy (2008) suggests winning games it the best way to maximize revenue.[1] Thus, this paper makes the simplifying assumption that teams and general managers (the executive responsible for managing revenue and cost, a type of CEO) will focus only on winning games when making contract decisions. Under this assumption, general managers are expected to make decisions to offer contracts to players based on the future production they expect that player to bring to their team as they need to optimize their team payroll in order to gather the best collection of talent possible. The salary cap, a limit on how much money can be allocated to the players on a team, serves as a budget constraint on which to optimize payroll allocation.[2] These contract determination decisions are made using past performance data to project the future value of a player in terms of performance.

Due to the high stakes nature of NBA contract decisions, precise statistical records are kept and analyzed in order to determine which players can give each team the best chance of winning basketball games. Using this data, general managers should be able to predict player performance accurately, causing them to be paid in accordance with their future performance. However, it is likely that some of the aforementioned psychological biases will complicate and alter the general managers' decision-making processes, causing them to overvalue performance

[1] While there is an argument that some NBA teams intentionally lose in a given year in order to receive higher draft picks, thereby potentially increasing their chances of winning the following season, that will be ignored for the purposes of this paper.

[2] The assumption of this paper is that teams will spend up to the limit allowed under the luxury tax system if it will increase their chances of winning games. So, the salary cap can be treated as a limit that all teams will hit in trying to win games if they are capable of being competitive and the marginal tradeoff of dollars per win is appropriate.

during the "contract year" (the year prior to free agency).[3] This paper will investigate whether NBA general managers display rationality in contract determination decisions in order to provide evidence for systemic irrationality in performance pay labor markets. The hypothesis of this study is that NBA general managers will overweight performance during the contract year at the expense of performance from years prior to contract year, providing evidence that psychological biases inhibit the processes of rational decision making.

The overweighting of the previous year's performance is exacerbated by the "contract year effect." The contract year effect is the concept that a player will elevate their level of play in an attempt to garner a longer term or more highly valued contract due to perceived incentives for good performance. A possible manifestation that causes the contract year effect is the willingness to play through injuries in the contract year as many studies have linked contract value to cumulative statistics that increase by playing more. After the contract year has passed, the performance of these players often returns to previous levels due to a number of factors including regression to the mean, less focus, or less desire to play through injury.[4] A counteracting theory to irrationality caused by the contract year effect is the idea of non-stationarity among player performance in the NBA. Non-stationarity is the idea that a player can improve or get worse over the course of his career. If this is true, which is a common belief in the sports world, then contract year performance should be a better indicator of his current ability and thus a better predictor of his future performance than years prior to contract year. Due to these factors, if there was uncharacteristic performance in a contract year, the myriad of

[3] A general manager will also suffer from the status quo bias, documented in Kahneman et. al (1991), whereby the disadvantages to leaving the status quo loom larger than the advantages, causing general managers to continue to value players the same way the rest of their peers do, even if it shows signs of irrationality.
[4] Skeptics of the contract year effect please note: the contract year effect is closely tied to a concept found in the private equity market by Chung et. al (2010) whereby rational "agents will take actions to maximize people's perception of their abilities" even when it is not in the best interest of the firm to do so thereby increasing their own bonuses.

confounding cognitive biases and beliefs may cause general managers to accept rather than reject that performance as indicative of future performance, which is irrational as performance always tends to regress to the league mean and career average levels of production.

In a movement begun in baseball known as sabermetric analysis, basketball statisticians have increased their desire and ability to capture player performance as precisely as possible. More accurate data is kept, analyzed, discussed, and compared across players. While this data is valuable for analyzing player performance, it also provides a proxy by which to capture executives' performance by computing their ability to efficiently allocate capital in order to earn team wins. The advanced statistic win share is the catch-all statistic that will be used in this analysis and credits players for the portions of a win that they contribute during each game. For example, Kevin Durant led the NBA last year with a win share of 19.2 meaning he contributed play that accounted for 19.2 of the Oklahoma City Thunder's wins.[5] This statistic is chosen because it provides a full picture of all of the ways in which a player contributes to winning a basketball game (including non-standard or commonly cited statistics) which should be the exact criterion upon which players are paid. More traditional statistics such as points per game and assists per game fail to capture many defensive metrics such as how often the player an individual is guarding scores, which are captured in win share.

According to economic research, the more clear, persistent, and straight forward performance feedback is, the more rationally it can be evaluated and utilized in decision making. Sports provide very clear and consistent data as well as a transparent market in which to investigate performance pay decision making due to the proliferation of available performance

[5] For a longer description of win share, please see the appendix.

data and the fact that player contracts are public knowledge. Thus, sports are the ideal market in which to investigate rational decision making and information processing based on performance pay. Further, basketball provides a setting where irrationality is as unlikely to occur as in any other setting due to year to year variation in performance being lower than it is for all other professional sports (Berri et. al, 2007). As a result the performance information for NBA players is as clear, consistent, and persistent as it can be in any labor market which should lead to rational expectations of future performance. While the contract year performance should receive a large portion of contract consideration as it is most closely tied to future performance, a rational decision maker is expected to accurately weight all prior performance data in contract determinations. Thus, while it is unlikely that irrationality would be present in NBA free agent contract determination decisions, any evidence of contract misappropriations would provide very strong evidence of the power and magnitude of cognitive biases and their effect on decision making.

The analysis in this paper will attempt to answer whether general managers of professional basketball teams pay players in a manner that exhibits recency bias. The answer to this question will give insight into the ability of managers in other labor markets to exhibit rationality and fairly compensate their employees. The analysis will be completed using two slightly different multi-step approaches. The first approach will look at the performance of a player surrounding their contract and the resulting contract terms received by a player. For methodology A, the first step will be to run a model that tests the impact of lagged performance on future performance, average yearly contract value, and length of contract simultaneously. In this model, the performance in the two years after receiving a contract are used as a proxy for future performance. The second step will be to compare the coefficients given to the lagged performance values in the future performance model to those in the contract models to determine if the lagged performance variables are adequately weighted in the contract model given their predictive power of performance.

For methodology B, the first step will be to run a model that tests the impact of lagged performance on future performance and total contract value simultaneously. In this model, the future performance is measured as the performance output of a player over the duration of their contract. The second step will be to compare the coefficients given to the lagged performance variables between the two estimation models to determine if the lagged performance variables are adequately weighted in the contract model given their predictive power of performance over the course of the contract. The coefficients in both methodologies will be compared using cross equation ratios that will demonstrate whether the lagged performance coefficients are being weighted with equality in contract and future performance models, the metric for rationality in this paper.

Literature Review

Most of the research concerning contract rationality has been done in Major League Baseball due to the early trend towards advanced statistics and appropriate talent evaluation in the MLB. Most of these papers agree that contracts are determined with systemic inefficiency in the use of performance data. Each paper attributes this effect to different reasons but the well-documented effects of many cognitive biases provide a strong case for many of the inefficient confounds.

Gilovich et. al (1985) conduct a study on the hot hand fallacy in basketball in three different studies and found that people tend to systemically correlate recent results with the next result. This provides evidence of both recency bias and a belief in non-stationarity in shooting talent over short periods of time, an effect that is a common belief about all sports performance. This belief in non-stationarity is due to people's biased beliefs about probability, as psychological studies have proven that people often assume there will be more alternation in random patterns than statistics predict. In a different study, Colin Camerer (1989) looked for the belief in positive recency, or recency bias, in the sports betting market and found that betting against teams on "a winning streak" or for teams on a "losing streak" can lead to beating the sports betting market. Again, this is evidence the well-documented hot-hand fallacy and stems from recency bias and people's inability to detect randomness in ordered sequences of outcomes. These studies point to the inability of most individuals to de-bias their decision making processes when it comes to evaluating sports performance.

Andrew Mooney (2011) does an analysis of large free agent contracts in Major League Baseball. He examines wins above replacement (the baseball statistic comparable to win shares) of free agents who received large contracts. He finds that almost no player achieved their level of past performance during their free agent contract. In spite of this, free agents were paid based on the assumption that their contract year performance would be predictive of their performance during their contract. The systemic overpayment is relevant to this paper because it was found that the performance of those free agents during their contract was closer to matching their performance over the multiple year period prior to the contract year than in the year just preceding the receipt of a contract.

Hochberg (2011) analyzes how contract year performance affects salary in his paper, "The Effect of Contract Year Performance on Free Agent Salary in Major League Baseball." He uses data from position players between the 1993 and 2010 seasons to look at the effect of contract year performance on contract determination. He investigates whether contract year performance is a better indicator of future performance or of the free agent contract. Hochberg finds that, "Offensive production in period (t – 3) may be underweighted and offensive production in a contract year (t – 1) may be over-weighted when determining salary relative to its effect on performance in year t," (22). He points to positive recency and a belief in non-stationarity as possible explanations for misallocation of contracts to free agents, finding documented evidence of contract determination irrationality in Major League Baseball. However, it has been found that consistent, clear, and persistent feedback should negate the effect of these types of psychological biases (Camerer and Lovallo, 1999). An investigation for these same tendencies in the NBA would provide stronger evidence of the power of psychological biases as it is a more transparent job market than even the MLB. Therefore, any evidence of psychological biases causing oversight of past performance would imply the power and strength these biases have in everyday decisions (including contract determinations in all job markets) due to performance in the following season being so predictable in the National Basketball Association given the lagged performance data.

Healy (2008) conducts a similar analysis to Hochberg. He looks at data from 1985 to 2004 and runs models to determine the predictive power of past performance on future performance. He finds that many teams underutilize information from years prior to contract year in projecting future performance and thus often overpay players relative to their expected performance to a large extent. He also found that teams that utilize past performance data better win more games. This finding could generalize to the job market by stating that firms who are more efficient with their payroll by allocating based on all performance data available for their employees will earn higher profits. Healy also points to psychological biases for this inefficiency and posits some hypotheses to explain the "contract year effect." Healy points to the availability heuristic from Tversky and Kahneman's "Availability: A Heuristic for Judging Frequency and Probability", stating that general managers are considering only information that is easily recallable such as last year's performance to explain the misappropriation of contracts. Due to the correlation between player performances from year to year being the highest in the NBA of all professional sports, finding evidence for cognitive biases causing inefficient decision making would be even more revealing due to the relative ease of projecting future performance. The implications across all job markets are quite large as salaries are calculated in many job markets based upon performance with much less feedback in the form of performance statistics and with much less year to year performance correlation causing a potential for very large salary and bonus misappropriations.

These two studies, Hochberg and Healy, capture much of the effect in the MLB that this study hopes to find in the NBA. Berri et al. (2007) provide evidence that player performance is much less variable year to year in the NBA than it is in the MLB. Thus, an investigation into this topic in the NBA should add to the relevant literature on recency bias and the availability heuristic in the job market by examining a market where it is even less rational that these biases exist. Due to the high stakes nature of negotiating these contracts, one would expect that general managers would de-bias their decision making process and appropriately utilize all of the available information rather than relying on heuristics that suffer from cognitive biases. Yet, evidence has shown that NBA general managers may fail to exhibit rationality in their contract determination decisions.

Recency bias and other psychological effects are present in NBA decision making as Berri et. al (2007) state that “decision making in sports has been shown to be inconsistent with the precepts of instrumental rationality” in their paper on NBA contract determination. They claim that the contract allocation inefficiency is quite dramatic as positive recency and other cognitive biases are magnified by the “contract year effect.” As a result, much of the information that is salient to rational contract determination will be ignored and systemically underutilized in contract determination. As Barnes et. al (2012) found in their paper, both performance mean and trend have a strong positive relationship with change in compensation levels for NBA players. While performance mean is accurately weighted, they point to the fact that performance trend is systemically overweighted as the reason for consistent overpayment of free agents in the NBA given their future performance. Evidence of this systemic inefficiency would be revelatory because it would be occurring in a setting that provides the best chance for bounded rationality and Bayesian updating of beliefs due to the availability of performance data, the high stakes nature of the decisions, and the strong predictive power of past performance data. So, in addition to adding a documentation of these confounding psychological biases in another environment, this paper attempts to provide evidence for irrational decision making that is inconsistent with rational expectations of the future in an environment in which irrationality is as unlikely to exist as it is in any performance pay market.

Data

The data from this paper came from two different sources. The data for free agent contracts is taken from spotrac.com, a professional sports contract data website. From this source, all of the free agent contracts between the 2006-2007 season and the 2011-2012 seasons are taken.[6] This list only includes full season contracts that were signed by free agent players, as shorter length contracts are intended for a subset of players who are not comparable to the rest of the players in the NBA due to lack of data by which to measure performance. This paper employs contract data from the 2007 to 2012 time range as there is no data available on player contracts prior to the 2007 season and the impact on future performance cannot be determined for more recent free agents than the 2012 season. The contract data employed for this study are total dollar value of a contract and contract duration.

The data for player performance statistics are taken from basketball-reference.com. The performance data taken from this website span the seasons between 2005 and 2014 and is grouped by performance during each season. The performance statistics that have been utilized for this analysis are win share, minutes played, age, and team. All other performance statistics have been omitted as they are captured by win share, a catch-all statistic that quantifies all of the contributions a player makes towards his team winning a game in units of portions of a team win. Minutes played will ensure that the players are compared on an even playing field so that there is no issue of cumulative win share driving up contracts simply via more playing time. Age is included because average production tends to decline with age. Team during the contract year is

[6] Each season will be called by the ending season year from here onwards, e.g. the 2006-2007 season will be called the 2007 season. This also coincides with the term contract year as a player whose contract year occurs in the 2006-2007 season will likely sign their free agent contract in the summer of 2007.

included as each team may have its own spending strategy based on size of market and how close they are to being able to compete for a championship in a given season.

The production variable of interest in this study will be win share. Win share attempts to credit a player with portions of a contribution to a team win during a specific game. These values can be positive or negative and reflect a composite of all of the different ways in which a player impacts the game on both offense and defense. A more detailed description of how win share is calculated is included in the appendix. The mean win share during the contract year for all players in the data set is 2.69 wins while all of the players had a win share value between -1.3 and 18.5 wins. The lead and lagged win share values for each contract vary slightly but are similar to these values. In addition to this performance variable, the control variables utilized are age and minutes played with team and year dummy variables.

In order to determine whether contract determination decisions were made rationally, lead and lagged performance variables were attached to each free agent contract. Thus, each contract has a value for win share, minutes played, and age for contract year, two years prior to contract year, one year prior to contract year, one year after contract year, and two years after contract year. The natural log of win share, contract duration, and dollar value of contract is used to avoid a right hand skew.[7] The final data set includes 367 different contract observations for analysis.

Each position in basketball requires a unique set of skills. As a result, player salaries are determined at least in part by the skill set a position requires as well as the rarity of the skills or competency of the skill set a specific player has. For example, it is much more difficult to find a

[7] Because the statistic win share has values ranging as low as -1.7, all win share values had two wins added to them when the natural log was taken to keep the numbers feasible.

skilled center (a position that requires a player to be at least 6’10”) than a skilled point guard who can be any height (taller than 5’10”). This data set has 81 point guards, 65 shooting guards, 81 small forwards, 77 power forwards and 63 centers. However, according to the literature, only centers are paid in a different manner than the other four position players as they are the only player whose skill set cannot be approximated adequately using a player who is not above 6’10”.

There will be two main methodologies for investigating the rationality of contract determination in this paper. The first will use performance in the two years prior, the contract season, and the two seasons following the receipt of a contract to investigate the rationality of contract determination decisions. It will investigate the rationality of the average contract value the players are paid on a yearly basis and the length of the contract they are given simultaneously and independently. The average performance in the two years after receiving a contract will be used as the proxy measure for future production as it allows the rationality of performance forecasting on the part of general managers to be tested under a stable and certain time horizon. One season is unstable due to random variation causing uncertainty and more than two seasons is unstable due to aging effects or injury possibilities.[8]

Table 1 details the main dependent and explanatory variables of methodology A. The summary statistics are calculated only for the contracts whose players have all five years of performance data surrounding their free agent contract as those are the contracts of interest for this paper. The average free agent contract is $19 million with an average yearly contract of $5.15 million over 2.72 years.

[8] The rationality of paying players beyond the two years after receiving a contract will be left out of this methodology, but will be captured in methodology B.

Table 1: Summary Statistics of Main Dependent and Explanatory Variables

Variables	Mean	St. Dev.	Min	Max
Dollars	19,000,000	24,200,000	540,872	118,000,000
Avg Yr Con	5,157,755	4,356,660	540,872	19,900,000
Duration	2.72	1.61	1	6
WS2CY	3.39	2.81	-0.7	18.5
WS1CY	3.48	3.10	-0.6	15.2
WSCY	3.44	3.02	-1.2	20.3
WSCY1	3.23	2.78	-1.5	15.6
WSCY2	2.99	2.74	-0.7	14.5
Age	27.16	3.64	20	38

*****Note WS2CY refers to win share two seasons prior to contract year, etc. Age refers to a player's age during their contract year.

Production tends to decline as time passes due to aging effects causing physical skill deterioration as shown by the mean of win shares decreasing from one year prior to contract year to two seasons after contract year. Further, basketball players peak in performance at age 27 according to Paine (2009) which is the average age of the players in this study and coincides with the average decline in performance after receiving a contract.

The contract variables of interest in this study are total contract dollars, duration, and average yearly contract. Table 2 details the summary statistics for methodology A by contract length (in years):

Table 2: Summary Statistics by Contract Length

Duration	Contract Value	Avg Yearly Contract	Age	WS2CY	WS1CY	WSCY	WSCY1	WSCY2
1	2,191,054	2,191,054	28.14	2.34	2.08	1.77	1.99	1.86
2	6,995,462	3,497,731	27.19	2.95	2.84	2.66	2.65	2.52
3	14,900,000	4,952,709	27.86	3.58	3.46	3.26	3.22	3.01
4	31,000,000	7,752,409	26.11	3.89	4.38	4.59	4.41	4.03
5	46,500,000	9,302,912	25.83	5.55	4.90	5.59	4.54	4.31
6	75,800,000	12,600,000	25.47	6.02	7.51	7.65	6.47	5.43

As expected, the average yearly contract value and average win share (for all seasons) increases as duration increases, implying that better players warrant and sign longer contracts than those who receive shorter contracts. Average age also declines as contract length increases which abides by the conventional logic that production will decline with age. There is mild evidence for the contract year effect amongst those players who receive contracts longer than three years in length (who also tend to be better players) as it appears as though the free agents elevate their play during contract year and return to normal levels of production thereafter. This data could provide evidence that those players who are highly talented can manipulate their performance output during the contract year in an attempt to secure a more highly paid long term contract, which is further reinforced by the decrease in production the year after receiving a contract.

The second methodology of the paper, methodology B, will investigate the rationality of a contract decision made by a general manager. It will match all of the performance data a player accumulates throughout their contract and compare it to the total amount of money they will receive under their contract. Therefore, it will determine the total rationality of the contract given the available lagged performance data for a player. Table 3 details the main descriptive statistics of methodology B organized by length of contract:

Table 3: Summary Statistics for Methodology B by Contract Length

Years	Dollars	WS2CY	WS1CY	WSCY	Tot Fut WS	Age
1	1,922,908	2.17	1.89	1.49	0.16	28.60
2	6,995,462	2.95	2.84	2.66	2.71	27.19
3	13,800,000	2.79	2.92	3.25	6.21	26.79
4	26,200,000	3.68	4.80	4.98	12.15	26.47
5	43,900,000	5.31	5.44	5.66	17.51	25.64
6	63,500,000	3.76	5.49	6.03	23.13	24.43

Total future win share is the sum of all of the win share values a player accumulates throughout the length of their contract. Age tends to decrease as length of contract increases. Past production and future production over the course of the contract increase as contract length increases which validates the idea that better players receive longer contracts. There is also evidence for the contract year effect amongst the more talented players who received longer contracts. These players seem to elevate their level of play during the contract year as the win share value during contract year is higher than in the years prior to contract year and the average level of production over the course of a contract is well below the contract year level. The fact that players production tends to decline on average from contract year to years under contract provides mild evidence that NBA players may be overpaid if they are paid on the basis of their most recent performance only.

Model/Methodology

In order to investigate whether general managers suffer from recency bias or the availability heuristic in contract determination decisions, the effect of lagged performance on future performance and contract value must be determined. This will be determined using two distinct methodologies. Methodology A will focus on how past performance affects future performance in the two years after receiving a contract in order to investigate contract determination decisions over a period which should be predictable and relatively invariant. However, it will not investigate whether long term contracts obey the tenets of strict rationality given the available past performance data used for projecting future performance. Methodology B will investigate the rationality behind free agent contracts in their entirety, matching the future performance of a player during their contract to the amount of money they are paid over the duration of their contract.

Methodology A

The first model run in order to determine the rationality of free agent contracts in the NBA will have three main estimation models. The model is a multivariable seemingly unrelated regression, which estimates three separate equations simultaneously and allows for correlated error terms. The first equation will estimate how lagged performance affects future performance. The second and third equation will determine how past performance affects both average yearly contract value and the length of a free agent contract. The three equations being estimated are detailed here:

$$LnFutureWinShares = \beta_0 + \beta_1(LnWinShare_{CY}) + \beta_2(LnWinShare_{CY-1}) + \beta_3(LnWinShare_{CY-2}) + \beta_4 (Abs_{cy-1}) + \beta_5(Abs_{cy-2}) + \beta_6(Age_{CY}) + \beta_7(Age^2_{CY}) + \beta_8(C) + \beta_9(MP_{cy}) + \beta_{10}(MP_{cy-1}) + \beta_{11}(MP_{cy-2}) + e_i$$

$$LnAverageYearlyContractValue = \beta_0 + \beta_1(LnWinShare_{CY}) + \beta_2(LnWinShare_{CY-1}) + \beta_3(LnWinShare_{CY-2}) + \beta_4 (Abs_{cy-1}) + \beta_5(Abs_{cy-2}) + \beta_6(Age_{CY}) + \beta_7(Age^2_{CY}) + \beta_8(C) + \beta_9(MP_{cy}) + \beta_{10}(MP_{cy-1}) + \beta_{11}(MP_{cy-2}) + e_i$$

$$LnDuration = \beta_0 + \beta_1(LnWinShare_{CY}) + \beta_2(LnWinShare_{CY-1}) + \beta_3(LnWinShare_{CY-2}) + \beta_4 (Abs_{cy-1}) + \beta_5(Abs_{cy-2}) + \beta_6(Age_{CY}) + \beta_7(Age^2_{CY}) + \beta_8(C) + \beta_9(MP_{cy}) + \beta_{10}(MP_{cy-1}) + \beta_{11}(MP_{cy-2}) + e_i$$

The first equation investigates how previous performance models future performance. The natural log of future win shares is the natural log form of the average win share a player accumulates over the two years after receiving a contract plus two wins. Two wins are added to win share values of all seasons prior to taking the natural log so that all win share values will be positive. The main explanatory variables of interest are the natural log of win share during the contract year, one year prior to contract year, and two years prior to contract year. The dummy variables for whether a player was present during a season prior to contract year were added so that players who were absent during one of the years preceding contract year could be included in the regression with a 0 value in place of a missing value for that lagged performance variable. The variable was given a value of 1 if a player was absent during a particular season. The dummy variable for the position center was included as Berri et. al (2007) found there to be a significant difference in pay level for centers due to their relative scarcity which is reflected in the data. Additionally, age and age squared were included in an attempt to capture the impact that a player's age has on their future production as most basketball players peak in performance around age 27 according to Paine (2009) which is the mean of age during the contract year in this study. Minutes played during each of the seasons of analysis is controlled for so that players who play more are not given an unfair advantage simply by accumulating win shares during their extra playing time.

The first equation specifies the extent to which past performance affects future performance. The second and third equations are run simultaneously with the first equation to investigate the impact of past performance on free agent contract terms. Because free agent contracts have two key components, length of contract and total dollar amount of contract, two separate equations for contract were included in the seemingly unrelated regression model.

The win share values for contract year (CY), one year prior to contract year (CY-1), and two years prior to contract year (CY-2) are the coefficients of interest. The coefficients on win share during each of the three seasons prior to receiving a contract will show the size of the effect an increase in win shares in each of the years prior to receiving a contract has on future performance, average yearly contract value, and duration of contract. As mentioned previously, the win share statistic is designed to take into account all positive and negative contributions to the team on the offensive and defensive sides of the game. The same control variables are used in all three equations as the assumption of the paper is that general managers pay players based on their projection of the player's future performance.[9] All three equations are estimated with and without team and year fixed effects to account for different spending strategies by different teams, to account for inflation, and to account for league wide increases in the salary cap, while still giving insight into the explanatory power of the independent variables.

[9] Note: While Hochberg (2011) points to a study by Kahn that implies that total contract value should be used, the twofold approach taken to capture contract value should capture the nature of the contracts more precisely.

Methodology B

Methodology B will investigate the rationality behind the contract a free agent receives during their contract year. It will compare the contract they receive in total dollar amount to the total number of win shares they contribute to their team over the course of their contract. For this analysis, two equations will be estimated simultaneously using a seemingly unrelated regression which allows for the two equations to have correlated error terms. Similar to methodology A, the equations for future performance and contract terms will be run simultaneously to investigate the impact lagged performance has on future performance and on free agent contracts. The equations will be estimated using the following forms:

$$LnTotalFutureWinShares = \beta_0 + \beta_1(LnWinShare_{CY}) + \beta_2(LnWinShare_{CY\text{-}1}) + \beta_3(LnWinShare_{CY\text{-}2}) + \beta_4(Abs_{cy\text{-}1}) + \beta_5(Abs_{cy\text{-}2}) + \beta_6(Age_{CY}) + \beta_7(Age^2{}_{CY}) + \beta_8(C) + \beta_9(MP_{cy}) + \beta_{10}(MP_{cy\text{-}1}) + \beta_{11}(MP_{cy\text{-}2}) + e_i$$

$$LnDollars = \beta_0 + \beta_1(LnWinShare_{CY}) + \beta_2(LnWinShare_{CY\text{-}1}) + \beta_3(LnWinShare_{CY\text{-}2}) + \beta_4(Abs_{cy\text{-}1}) + \beta_5(Abs_{cy\text{-}2}) + \beta_6(Age_{CY}) + \beta_7(Age^2{}_{CY}) + \beta_8(C) + \beta_9(MP_{cy}) + \beta_{10}(MP_{cy\text{-}1}) + \beta_{11}(MP_{cy\text{-}2}) + e_i$$

The dependent variable in the first equation is the natural log of total future win share is the sum of the natural log of win share values that a player accumulates over the course of their contract. The main coefficients of interest are the coefficients on the lagged performance variables for the natural log of win share during the contract year, one year prior, and two years prior. Again, dummy variables for whether a player was absent from the NBA for one of the seasons prior to contract year was included to allow for these players to be included in the regression. The dummy variable for the position center was included as centers are paid in a different manner than all other position players. Age and age squared were included in an attempt to capture the impact that a player’s age had on their future production. Minutes played during each of the seasons prior to receiving a contract is controlled for so that players are rewarded for their marginal production rather than their total production. This model is also run with and without team and year fixed effects for reasons specified earlier.

The second equation is an investigation into how past performance affects the contract a free agent receives in total dollar value. The natural log of dollars is the natural log of the total dollars a player receives as part of their free agent contract assuming they are eligible, healthy enough to play, and do not opt to retire. The control variables are the same as in the future performance equation of methodology B so that the two equations can be compared directly and because general managers should be attempting to model future performance in their contract allocation decisions. Again the model is estimated using team and year fixed effects and without using them. As in methodology A, the coefficients of interest are the coefficients on win share during the contract year, one year prior to contract year, and two years prior to contract year.

Finally, in order to examine whether past performance was given the appropriate weighting across the models in methodology A and methodology B, coefficient ratios are calculated and compared across equations. The coefficients of interest are ratios of the form:

Ratio 1: $(\beta_{wscy-1}) / (\beta_{wscy})$ Ratio 2: $(\beta_{wscy-2}) / (\beta_{wscy})$

After running the regression and determining the coefficients from the previous models in methodologies A and B, a ratio test calculates these coefficient ratios and compares them across the estimations for A and B respectively. In effect, these ratios provide an appropriate weighting of the two lagged performance coefficients given their predictive power of future performance compared to contract year performance's impact on future performance. These ratios are then compared to the same ratio in the contract models to investigate whether the lagged performance data are being appropriately utilized. If the ratios are larger in the performance equation than in the contract equation, this ratio test provides evidence for underweighting of performance data prior to contract year in contract decision making processes which could be explained by recency bias and other cognitive biases such as the availability heuristic. The added benefit of using these ratios, in addition to their weighting relative to contract year performance, is that they are unit-less and thus allow for the two models to be compared directly. The comparison of these two ratios across the three equations of methodology A and two equations of methodology B will serve as the primary foundation for determining strictly rational contract allocation decisions on the part of NBA general managers.

The models specified previously used the natural log form of average yearly contract, contract duration, and win shares. The variables all showed a strong positive skew, as a result, scatter plots of the variables of interest were investigated and the logarithmic-logarithmic regression model versions of the scatter plots had the strongest linear relationship. The residual scatter plots were then compared and appeared the most cloudlike in the logarithmic-logarithmic regression model. As a result, the logarithmic forms were used in the estimation models.

Analysis and Results

Methodology A

The model utilized in methodology A investigated the impact between lagged performance variables and future offensive output, average yearly contract value, and contract duration. The main coefficients of interest from the model with team and year fixed effects are detailed in the table below.

Table 4: Effect of Past Performance on Future Performance, Average Yearly Contract and Contract Length with Fixed Effects

	Natural Log of Future Win Share		Natural Log of Average Yearly Contract Value		Natural Log of Contract Duration	
Variable	**Coefficient**	**Std. Err.**	**Coefficient**	**Std. Err.**	**Coefficient**	**Std. Err.**
$LnWS_{cy}$	0.4340***	0.0859	0.8499***	0.1262	0.6039***	0.1179
$LnWS_{cy-1}$	0.1620*	0.0838	0.0195	0.1231	0.1141	0.1150
$LnWS_{cy-2}$	0.1949***	0.0761	0.0468	0.1118	-0.0456	0.1045
Abs_{cy-1}	0.1038	0.1017	-0.1157	0.1493	0.0179	0.1395
Abs_{cy-2}	0.2540***	0.0886	0.1448	0.1301	0.1299	0.1216
Age	0.0089	0.0593	-0.1448*	0.0871	-0.1747**	0.0814
Age^2	-0.0006	0.0010	0.0018	0.0015	0.0024*	0.0014
Center	0.0538	0.0494	0.4270***	0.0725	0.1417**	0.0678
MP_{cy}	0.0000	0.0001	0.0001	0.0001	0.0000	0.0001
MP_{cy-1}	0.0000	0.0001	0.0002**	0.0001	0.0001	0.0001
MP_{cy-2}	0.0000	0.0000	0.0003***	0.0001	0.0000	0.0001
Obs.	349		349		349	
R^2	0.5888		0.7439		0.5233	

* indicates statistical significance at the 90% confidence level
** indicates statistical significance at the 95% level
*** indicates statistical significance at the 99% level

The significant results for the natural log of win share values in the future performance model suggest that all three years of production prior to contract year affect future performance of NBA players. As expected, the year just prior to receiving a free agent contract is the best indicator of future performance both in magnitude and in significance. A 5 percent increase in win share during the contract year is associated with a 2.12 percent increase in future win share (the average of win share over the two years following the receipt of a free agent contract) ceteris paribus. A 5 percent increase in win share in the year to contract year is associated with a 0.79 percent increase in future win share, and a 5 percent increase in win share two years prior to contract year is associated with a 0.95 percent increase in future win share, holding all other variables equal. Being absent from the NBA one year prior to contract year does not have a significant impact on the natural log of future win share; however, being absent from the NBA two years prior to contract year is associated with about a 29% increase in future win shares. Age during the contract year, age squared, the center dummy, and the minutes played variables do not have a significant impact on the natural log of future win share.

In the average yearly contract value and contract duration models, the natural log of win share during the contract year was highly significant. A 5 percent increase in win share during the contract year is associated with a 4.15 percent increase in average yearly contract value and a 2.95 percent increase in contract duration ceteris paribus. The lagged performance coefficients are not significant in either the average yearly contract value or the contract duration model, meaning they may not be utilized in contract determination decisions by NBA general managers. Players who were absent one or two years prior to contract year were not treated differently in contract determination processes in either average yearly contract value or contract length. Age was significant and negative for both contract models as expected holding all other variables constant. Age squared was positive and significant in the contract duration model, implying its importance in contract length decisions but not the average yearly value considerations. Centers were rewarded with a 53.27 percent increase in average yearly contract value and a 15.22 percent increase in contract duration all else equal. Finally, while minutes played during the years prior to receiving a contract had no significant impact on contract length, minutes played in years prior to contract year had a significant and positive impact on average yearly contract value. These findings ideas suggest that basketball players may be paid for consistency as those players who have played more minutes in the NBA in years prior to contract year are paid more money than those who played fewer or no minutes in those seasons ceteris paribus.

The Adjusted R^2 values for these regressions are high, signaling the explanatory power of the variables utilized in this estimation. In order to demonstrate that the explanatory power is not just due to the team and year fixed effects, Table 5 relates the results of the seemingly unrelated regression model without team and year fixed effects:

Table 5: Effect of Past Performance on Future Performance, Average Yearly Contract and Contract Length without Fixed Effects

	Natural Log of Future Win Share		Natural Log of Average Yearly Contract Value		Natural Log of Contract Duration	
Variable	**Coefficient**	**Std. Err.**	**Coefficient**	**Std. Err.**	**Coefficient**	**Std. Err.**
$LnWS_{cy}$	0.4385***	0.0814	0.8455***	0.1217	0.5889***	0.1139
$LnWS_{cy-1}$	0.2020**	0.0847	0.0661	0.1267	0.1755	0.1185
$LnWS_{cy-2}$	0.1768**	0.0762	0.0498	0.1140	-0.0189	0.1066
Abs_{cy-1}	0.1239	0.1017	-0.0624	0.1521	0.0673	0.1423
Abs_{cy-2}	0.2450***	0.0899	0.1485	0.1345	0.1536	0.1258
Age	0.0176	0.0580	-0.1300	0.0868	-0.1862**	0.0812
Age^2	-0.0008	0.0010	0.0015	0.0015	0.0025*	0.0014
Center	0.0512	0.0506	0.3891***	0.0757	0.1226*	0.0708
MP_{cy}	0.0000	0.0001	0.0001	0.0001	0.0000	0.0001
MP_{cy-1}	0.0000	0.0001	0.0001*	0.0001	0.0000	0.0001
MP_{cy-2}	0.0000	0.0000	0.0003***	0.0001	0.0000	0.0001
Obs.	349		349		349	
R^2	0.5305		0.6962		0.4338	

* indicates statistical significance at the 90% confidence level
** indicates statistical significance at the 95% level
*** indicates statistical significance at the 99% level

Comparing these results to Table 4 indicates that while the team and year fixed effects do slightly increase the adjusted R^2 of the model, a large majority of the explanatory power comes from the strength of the main independent variables. Aside from age during the contract year season in the average yearly contract model which is lowly significant with fixed effects and insignificant without, all of the coefficients have the same significance and direction. Further, they are very similar in magnitude, with all coefficients across the two models being well within one standard deviation of each other exhibiting the strength of the explanatory variables of this methodology.

The goal of this paper is to determine whether NBA teams and general managers in particular follow a rational decision making process in making contract determination decisions for free agent players. There was evidence found that performance in years prior to contract year were predictive of future performance but were ignored in contract determination decisions. The best way to determine rationality is to compare two ratios across both the future performance and the contract models. The ratios are of the form:

$$\text{Ratio 1: } (\beta_{wscy\text{-}1}) / (\beta_{wscy}) \qquad \text{Ratio 2: } (\beta_{wscy\text{-}2}) / (\beta_{wscy})$$

The ratios signify what percentage strength the lagged performance coefficients have in relation to the contract year performance coefficient. A smaller ratio would indicate that the contract year performance is much more significant and a ratio closer to one would indicate relative equality between the two coefficients in predicting the regressand. The two ratios will be calculated in each of the three equation estimations separately and then they are tested for equality. If the ratios were equal across the three models, then this would provide evidence that performance prior to contract year is appropriately weighted in contract decision-making processes given its predictive power of future performance. The ratios are calculated and described here:

Table 6: Coefficient Ratios from the Model with Fixed Effects

Model	Ratio 1	Ratio 2
Future Performance	0.3733	0.4491
Average Contract Value	0.0229	0.0551
Contract Duration	0.1889	-0.0755

Both ratio 1 and ratio 2 are larger in the future performance model than in either of the contract model specifications. This relationship between the ratios is consistent with the hypothesis that NBA general managers overweight contract year performance relative to performance prior to contract year given the predictive power of performance prior to contract year on future performance. Ratio 1 is not statistically significantly different across the models even at the most lenient 90% confidence level. However, ratio 2 is significantly different between the future performance model and the contract duration model at the 90% confidence level and is nearly significant between the future performance and average contract value model with a p-value of 0.11. This test finds evidence for irrationality on the part of NBA general managers in under-utilizing performance two years prior to contract year in contract length decisions.[10] While it appears that performance in years prior to contract year are ignored in large in part in contract determination despite having a strong predictive power of future performance, there is only mild evidence for irrational behavior in improperly weighting past performance data in contract determination decisions, particularly the data from two years prior to contract year.

[10] These findings are reinforced by the coefficient ratio tests in the seemingly unrelated regression model without fixed effects.

Methodology B

The model in Methodology B estimated the impact of lagged performance variables on future offensive output over the course of a contract and the total dollar amount a free agent receives. The main coefficients of interest from this fixed effects seemingly unrelated regression are detailed in the table below:

Table 7: The Effect of Lagged Performance on Total Future Win Share and Total Contract Value with Fixed Effects

	Natural Log of Total Future Win Shares		Natural Log of Total Contract Value	
Variable	**Coefficient**	**Std. Err.**	**Coefficient**	**Std. Err.**
$LnWS_{cy}$	2.2472***	0.464	1.3897***	0.2205
$LnWS_{cy-1}$	0.4142	0.4179	-0.0145	0.1985
$LnWS_{cy-2}$	0.3614	0.4233	0.0236	0.201
Abs_{cy-1}	0.6051	0.5023	-0.0774	0.2385
Abs_{cy-2}	0.6824	0.4601	0.1165	0.2185
Age	-0.4962	0.3175	-0.1984	0.1508
Age^2	0.007	0.0055	0.0026	0.0026
Center	0.1076	0.27153	0.4043***	0.1289
MP_{cy}	-0.0003	0.0003	-0.0001	0.0002
MP_{cy-2}	0.0000	0.0003	0.0001	0.0001
MP_{cy-1}	0.0001	0.0003	0.0004***	0.0001
Obs.	280		280	
R^2	0.6291		0.7129	

* indicates statistical significance at the 90% confidence level
** indicates statistical significance at the 95% level
*** indicates statistical significance at the 99% level

The results in Table 6 table indicate that the natural log of win share during the contract year has a significant and positive impact on future performance. A 5 percent increase in win share during the contract year has a 10.96 percent increase in total future win shares ceteris paribus. However, the lagged performance variables prior to contract year have no impact on future performance throughout the duration of an NBA contract. In fact, none of the other explanatory variables age, age squared, the dummy variable for the center position, minutes played for seasons prior to receiving a contract, or the dummy variables for being absent during a season prior to contract year had an effect on future production measured by the natural log of future win shares.

The natural log of win share during the contract year has a significant and positive effect on the total dollar value of a free agent contract model. A 5 percent increase in win share during the contract year season is associated with a 6.78 percent increase in the total dollar value of a free agent contract holding all other variables constant. However, like in the future production model of this methodology, the performance variables for years prior to contract year did not have a significant impact on total contract value. Centers were paid 49.83 percent more than the other position player free agents holding all else equal. The age and age squared control variables and the minutes played variables aside from one year prior to contract year did not have a significant impact on the total value of a free agent contract. Minutes played one year prior to contract year has a weak but positive effect on the total dollar value of a free agent contract, providing evidence that free agents may be paid for consistency of performing and playing at the NBA level. The adjusted R^2 values of the two equations are high, again signaling the strength of the regression models. In order to investigate the effects of the fixed effects on the explanatory power of the regression model, Table 8 was included displaying the output of the regression model without fixed effects:

Table 8: The Effect of Lagged Performance on Total Future Win Share and Total Contract Value without Fixed Effects

	Natural Log of Total Future Win Shares		Natural Log of Total Contract Value	
Variable	**Coefficient**	**Std. Err.**	**Coefficient**	**Std. Err.**
$LnWS_{cy}$	2.0934***	0.5087	1.3675***	0.2225
$LnWS_{cy-1}$	0.9335*	0.4944	0.1301	0.2163
$LnWS_{cy-2}$	0.1396	0.4844	-0.006	0.2119
Abs_{cy-1}	0.4879	0.5773	-0.1428	0.2526
Abs_{cy-2}	0.3825	0.5346	0.0415	0.2339
Age	-0.9185**	0.3621	-0.3371**	0.1584
Age^2	0.0135**	0.0064	0.0048*	0.0028
Center	0.1439	0.3228	0.4098***	0.1412
MP_{cy}	0.0002	0.0004	0.0001	0.0002
MP_{cy-2}	0.0001	0.0003	0.0001	0.0001
MP_{cy-1}	-0.0001	0.0003	0.0003*	0.0001
Obs.	280		280	
R^2	0.4248		0.6220	

* indicates statistical significance at the 90% confidence level
** indicates statistical significance at the 95% level
*** indicates statistical significance at the 99% level

The results of Table 8 demonstrate that while the fixed effects have an impact on the regression output of the seemingly unrelated regression from Methodology B, they do not drive the results. The adjusted R^2 value is still high and all of the coefficients are well within two standard deviations when compared across models. The main difference between the two models is that in the model without fixed effects, win share during the season one year prior to contract year becomes significant as a 5 percent increase in win share one year prior to contract year is associated with a 4.84 percent increase in total future win share accumulated over the course of a free agent contract holding all else equal. Additionally, the age and age squared control variables become significant in both the natural log of total future win share and the natural log of total contract value models holding all other variables constant. The age during the contract season variable is significant and negative while the age squared variable is positive and significant, as expected.

In order to test empirically for irrationality via inappropriate weighting of lagged performance coefficients, Ratios 1 and 2, described earlier, are compared across the two models of methodology B here:

Table 9: Coefficient Ratios in Methodology B with Fixed Effects

Model	**Ratio 1**	**Ratio 2**
Total Future Win Shares	0.1843	0.1608
Total Contract Value	-0.0104	0.017

While ratios 1 and 2 are bigger in the future performance model than in the total contract value model, showing evidence of the underweighting of lagged performance data, neither ratio is significantly different across the two models with any significance. Thus, there is no statistically significant evidence provided for irrationality on the part of NBA front offices and general managers in contract allocation decisions based on the performance data beyond one season prior to contract year. Methodology B provides mild evidence for underweighting of performance in years prior to contract year as both ratio 1 and 2 are bigger in the future performance model than in the total contract value model, potentially caused by recency bias and the availability heuristic, but there is no significant evidence for systemic irrationality by NBA general managers.

Conclusion and Discussion

The goal of this paper was to investigate the rationality of manager decisions in performance pay markets. The National Basketball Association's free agent market provides a strong opportunity to investigate this decision making process as there is an abundance of player (employee) performance data that is widely available, accurate, and thoroughly studied. Many documented psychological biases such as recency bias, the status quo bias, and the availability heuristic, may influence NBA front offices' decision making processes causing them to overweight contract year performance in contract decision making processes at the expense of the performance data from the two seasons prior to contract year.

Methodology A found that performance in seasons prior to contract year significantly predict the future performance of a player. However, it was found that performance one year and two years prior to contract year was largely ignored in the contract decision making process. Both one year prior to contract year and two years prior to contract year performance had no significant impact on either the average yearly dollar value of a contract or the length of the contract given. The coefficient ratio test from methodology A provided evidence for the underweighting of performance in years prior to contract year as both ratio 1 and ratio 2 were larger in the future performance model than the contract models, however there was no statistically significant evidence that performance one year prior to contract was systemically underutilized. However, there is evidence that performance during the season two years prior to contract year is overlooked by NBA general managers. The performance data from two years prior to contract year was significantly underweighted in contract duration decisions given its predictive power of future performance in the two years after contract year.

Methodology B did not find strong evidence that performance in years prior to contract year should be utilized in contract decisions as it was not found to be a significant predictor of performance over the duration of a free agent contract. There was mild evidence found that performance prior to contract year was underweighted by NBA front offices and general managers because ratios 1 and 2 were larger in the future performance model than the total contract value model but there was no statistically significant difference in comparing the ratios between models.

This study finds evidence for irrational behavior in the NBA free agent market in free agent contract decision making processes. Due to the nature of basketball, where variation in player performance from year to year is low, and professional sports, where performance feedback is clean, consistent, and widely known, any evidence of irrationality in performance pay decision making in this market is noteworthy. The NBA provides a market where it should be challenging to find evidence of irrationality, so any evidence found is significant in predicting a high potential for irrationality in other markets. The potential reasons for a failure to appropriately weight all performance data have been detailed at some length earlier but are beyond the scope of this paper, whether it be a belief in non-stationarity of player performance level, or one of a number of confounding psychological biases affecting what should be a rational, calculated decision using all available data. The results of this paper point to the strength of these psychological biases, mainly recency bias, as they appear to affect a decision which is expected to closely follow the tenets of strict rationality. There is evidence that performance data from two years prior to contract year is underweighted and improperly utilized in contract decision making processes and weak evidence that all performance prior to contract year is underutilized in contract determination processes. NBA front offices and general managers make contract determination decisions that ignore performance two years prior to contract year in ways that stray from strict rationality suggesting that such irrationality plagues decision makers in other performance pay markets where the performance data are noisier, more variant, and less widely known.

References

Arcidiacono, D. (2011). Consumer rationality in a multidisciplinary perspective. *Journal of Socio-Economics, 40*, 516-522.

Barnes, C., Ang, D., & Reb, J. (2012). More than just the mean: Moving to a dynamic view of performance-based competition. *Journal of Applied Psychology*, *97*(3),711-718.

Berri, David, Stacey Brook, and Martin Schmidt. "Does One Simply Need to Score to Score?."*International Journal of Sports Finance* 2007: 190-205. Web. <http://www.suu.edu/faculty/berri/ScoretoScoreIJSF.pdf>.

Calculating win shares. (n.d.). Retrieved November 24, 2014, from Basketball Reference website: http://www.basketball-reference.com/about/ws.html

Camerer, Colin F. "Does the Basketball Market Believe in the 'Hot Hand,'?" *American Economic Review*, 1989, 1257-61.

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *The American Economic Review*, *89*(1), 306-318.

Chung, J., Sensoy, B., Stern, L., Weisbach, M. (2010). Pay for performance from future fund flows: The case of private equity (Working Paper No. 16369). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w16369

Gilovich, Thomas, Vallone, Robert, and Tversky, Amos, "The Hot Hand in Basketball: On the Misperception of Random Sequences," *Cognitive Psychology*, July 1985, 295-314.

Healy, Andrew. "Do Firms Have Short Memories?: Evidence From Major League Baseball." *Journal of Sports Economics*: 407-424. Web. <http://jse.sagepub.com/content/9/4/407.full.pdf+html>.

Hochberg, Daniel. "The Effect of Contract Year Performance on Free Agent Salary in Major League Baseball." MA thesis. Haverford College, 2011. <http://triceratops.brynmawr.edu/dspace/bitstream/handle/10066/7050/2011HochbergD_thesis.pdf?sequence=1>.

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, New York: Farrar, Straus, and Giroux.

Kahneman, D., Knetsch, J., & Thaler, R. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, *5*(1), 193-206.

Kahneman, D., & Tversky, A. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, (5), 207-232.

Mooney, Andrew. "The Winner's Curse: Expensive Mistakes in Free Agency." *Boston.com*. The Boston Globe, Dec. 2011. Web. <http://www.boston.com/sports/blogs/statsdriven/2011/12/the_winners_curse_expensive_mi.html>.

Paine, N. (2009, July 19). Win shares and aging. Retrieved April 3, 2015, from Basketball-Reference.com website: http://www.basketball-reference.com/blog/?p=2953

Data Sources

Basketball-Reference.com. (2014). Individual Player Performance Data from 2004-2005 to 2013-2014. Retrieved from http://www.basketball-reference.com/leagues/

Spotrac.com. (2014). NBA Free Agent Tracker Data Sets 2006-2007 to 2011-2012. Retrieved from http://www.spotrac.com/free-agents/nba/

Appendix

The method for calculating win shares is detailed at length on BasketballReference.com on a page titled Calculating Win Shares found at: http://www.basketball-reference.com/about/ws.html.

Win Shares

One win is equivalent to one Win Share. If a basketball team were to have 55 wins, the sum of all of their players win share value would be 55 wins. A player can have negative win shares if they play so poorly that they actually hurt their team's chances of winning a game. A player's win share is made up of all plays both offensive and defensive that help or hurt their team's chances of winning a game.

Offensive win shares are calculated using some figures taken from Dean Oliver's book Basketball on Paper. First a calculation is made to determine how many points a player produces for their team and how many offensive possessions they had. From these two measures, marginal offense is calculated by scaling a player's contributions to the league average in points per possession. That marginal offense is then translated into win shares by dividing it by marginal points per win for a team in the NBA which takes into account pace of play of a player's team compared to the league and the average number of points per game.

Defensive win shares are also calculated relying on some of Dean Oliver's work. First, a player's defensive rating is calculated. Then their marginal defense is calculated which is scaled by minutes played, defensive possessions, and the expected number of points per possession in the league. Finally, defensive win shares are calculated by dividing marginal defense by the marginal points per win for a team which accounts for league pace, team pace, and points per possession

Finally, offensive and defensive win shares are summed to provide the total win share value for a player.

Data Appendix

Contract-Level Analysis Data Set

Methodology A

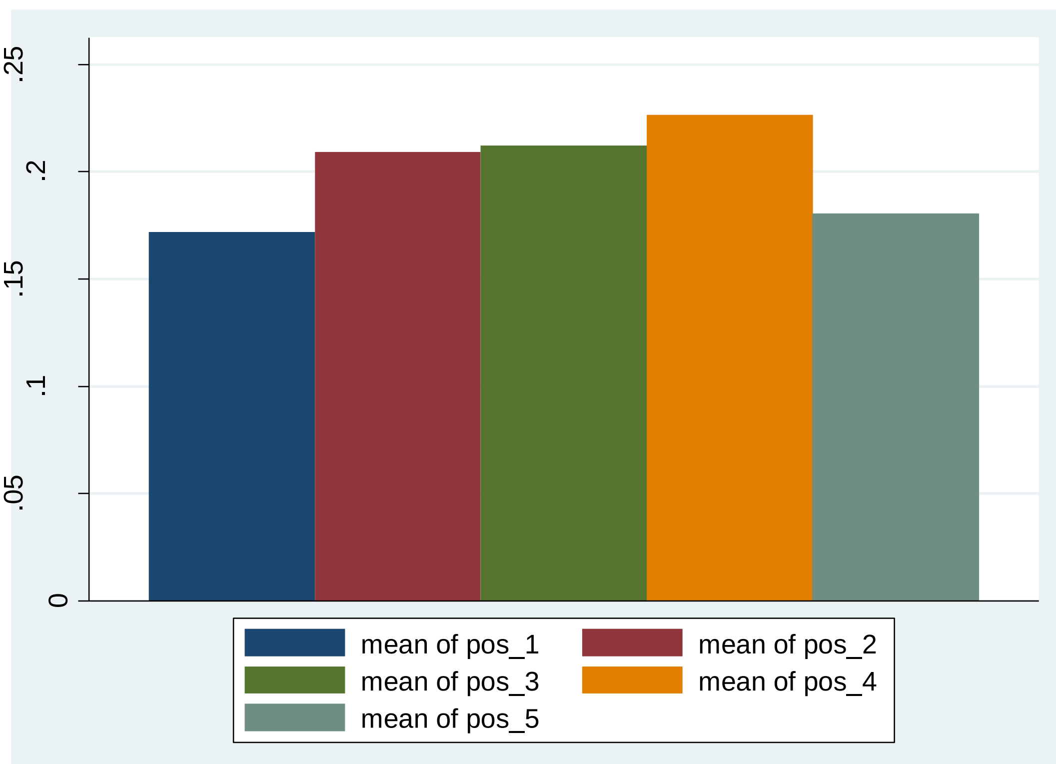
This data set includes all of the National Basketball Association contracts signed by free agents between the 2006-2007 and 2011-2012 seasons. There are 349 NBA contracts in this data set which contains information relating their performance across the three years prior and up to two years after receiving their contract. There are 248 unique players with 71 players appearing more than one time.

Variable name: *POS*
Source: Contract data set
Values: each of the five positions- Center, Power Forward, Point Guard, Small Forward, Shooting Guard

The distribution of contracts across positions is shown.

Pos	Freq.	Percent
C	60	17.19
PF	73	20.92
PG	74	21.20
SF	79	22.64
SG	63	18.05
Total	349	100.00

Frequency distribution of each position in bar graph form:

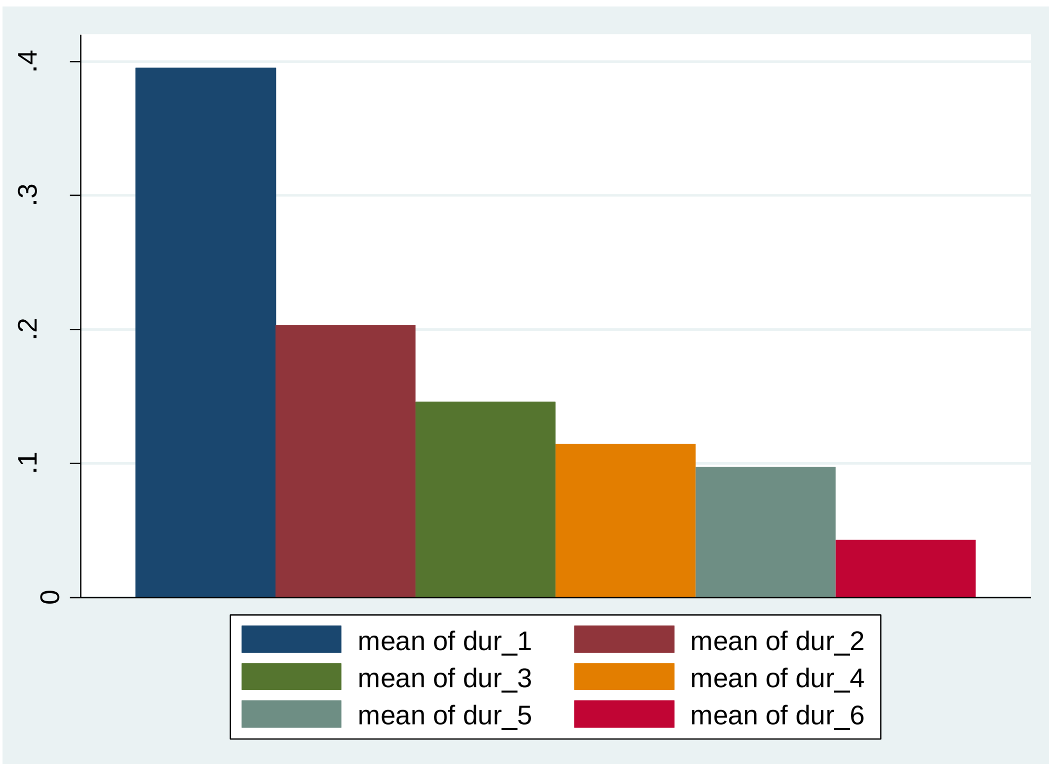


Variable name: *DURATION*
Source: Contract data set
Values: integers from 1 through 6
Coding: Length of contract in years

The distribution of contract durations is shown.

duration	Freq.	Percent	Cum.
1	138	39.54	39.54
2	71	20.34	59.89
3	51	14.61	74.50
4	40	11.46	85.96
5	34	9.74	95.70
6	15	4.30	100.00
Total	349	100.00	

Frequency distribution of each contract duration in bar graph form:

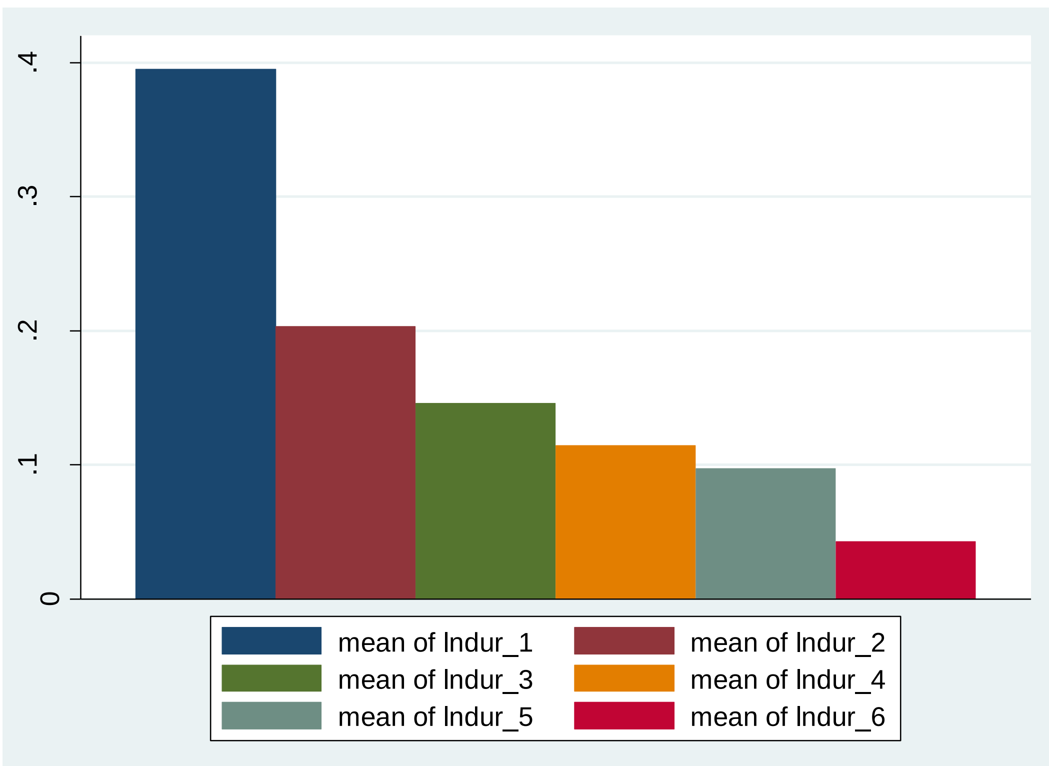


Variable name: *LnDURATION*
Source: Contract data set
Values: Numbers ranging from 0 to 1.8
Coding: Natural log of the length of contract in years

The distribution of the natural log of contract durations is shown.

lnduration	Freq.	Percent	Cum.
0	138	39.54	39.54
.6931472	71	20.34	59.89
1.098612	51	14.61	74.50
1.386294	40	11.46	85.96
1.609438	34	9.74	95.70
1.791759	15	4.30	100.00
Total	349	100.00	

Frequency distribution of the natural log of each contract duration in bar graph form:

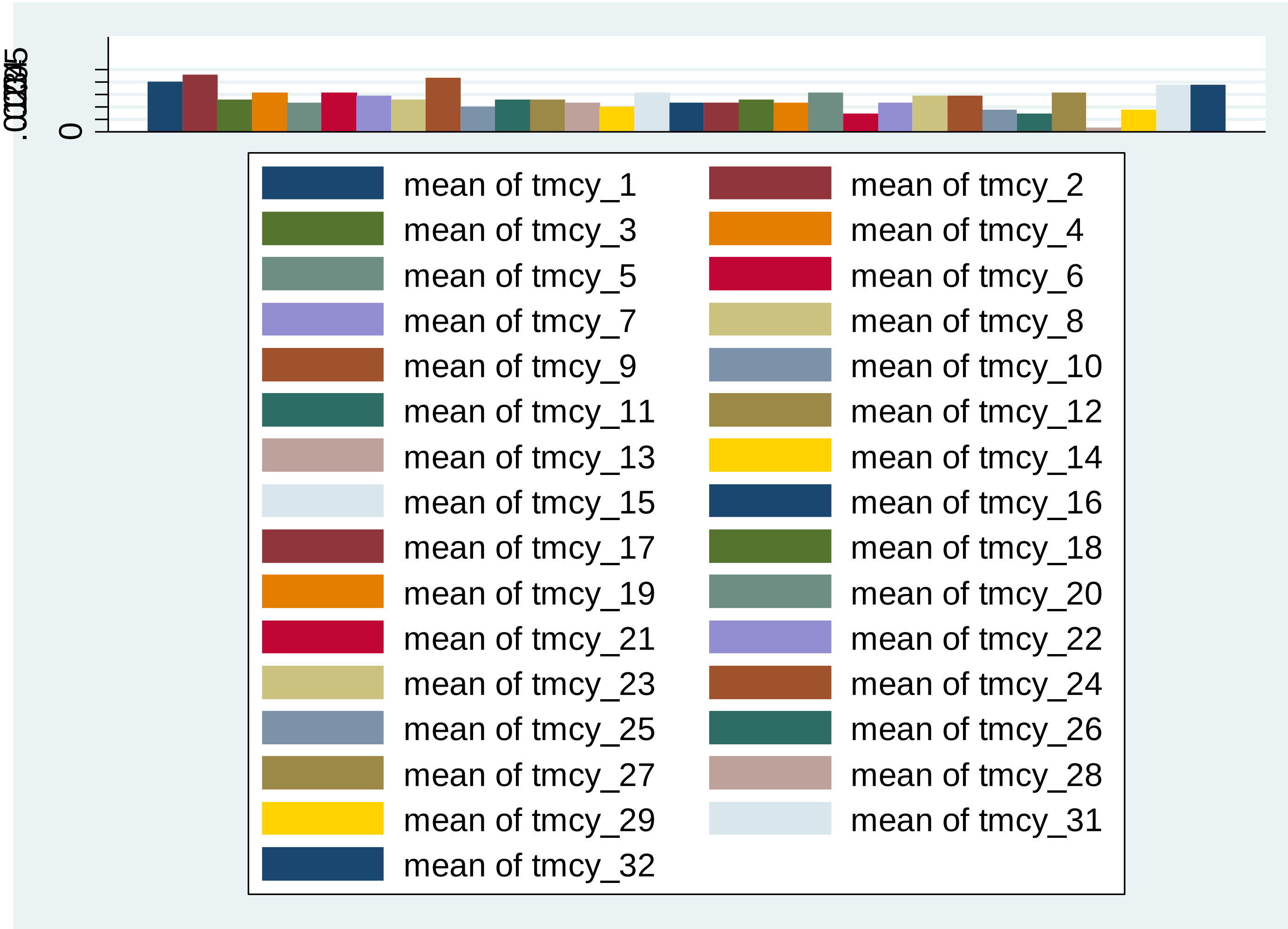


Variable name: *TMCY*
Source: Contract data set
Values: Each of the 31 different NBA teams over the period in question. *Players who switched teams during a season are not included here.

The distribution of contracts by team is shown:

Tm	Freq.	Percent
ATL	14	4.01
BOS	16	4.58
CHA	9	2.58
CHI	11	3.15
CLE	8	2.29
DAL	11	3.15
DEN	10	2.87
DET	9	2.58
GSW	15	4.30
HOU	7	2.01
IND	9	2.58
LAC	9	2.58
LAL	8	2.29
MEM	7	2.01
MIA	11	3.15
MIL	8	2.29
MIN	8	2.29
NJN	9	2.58
NOH	8	2.29
NYK	11	3.15
OKC	5	1.43
ORL	8	2.29
PHI	10	2.87
PHO	10	2.87
POR	6	1.72
SAC	5	1.43
SAS	11	3.15
SEA	1	0.29
TOR	6	1.72
UTA	13	3.72
WAS	13	3.72
Total	286	81.95

Frequency distribution of each team during contract year in bar graph form:

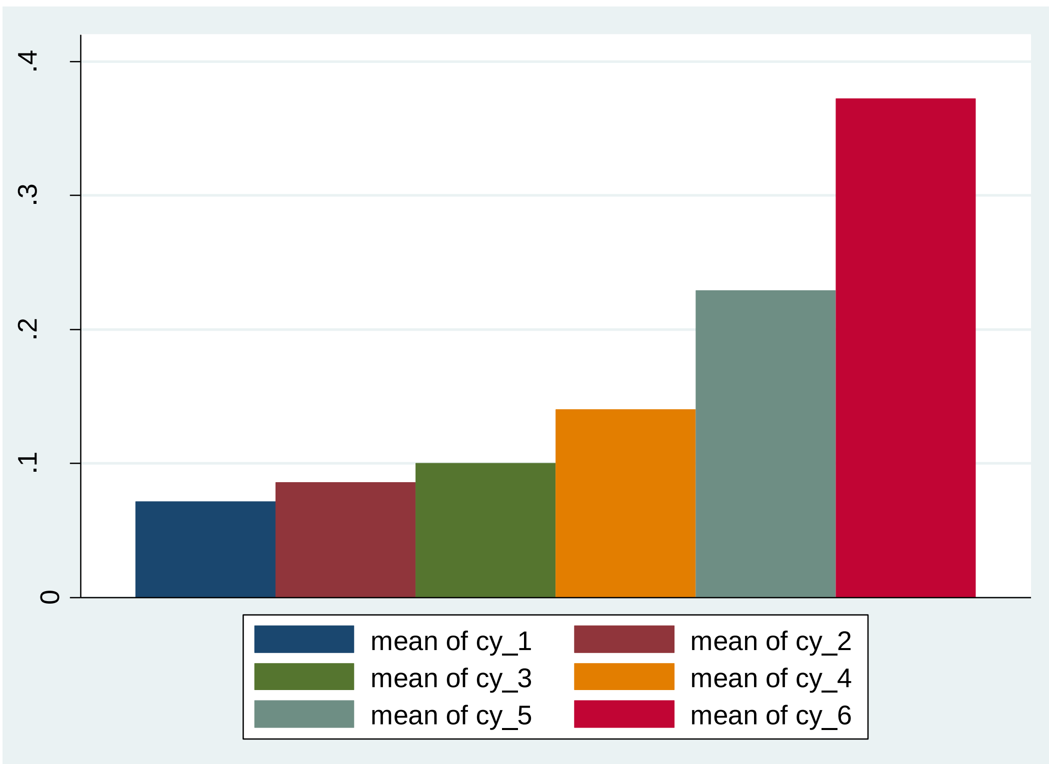


Variable name: *CY*
Source: Contract data set
Values: Integers ranging from 2007 to 2012
Coding: the year that a player signed his contract

The distribution of contracts by contract year is shown:

cy	Freq.	Percent	Cum.
2007	25	7.16	7.16
2008	30	8.60	15.76
2009	35	10.03	25.79
2010	49	14.04	39.83
2011	80	22.92	62.75
2012	130	37.25	100.00
Total	349	100.00	

Frequency distribution of contract by contract year in bar graph form:

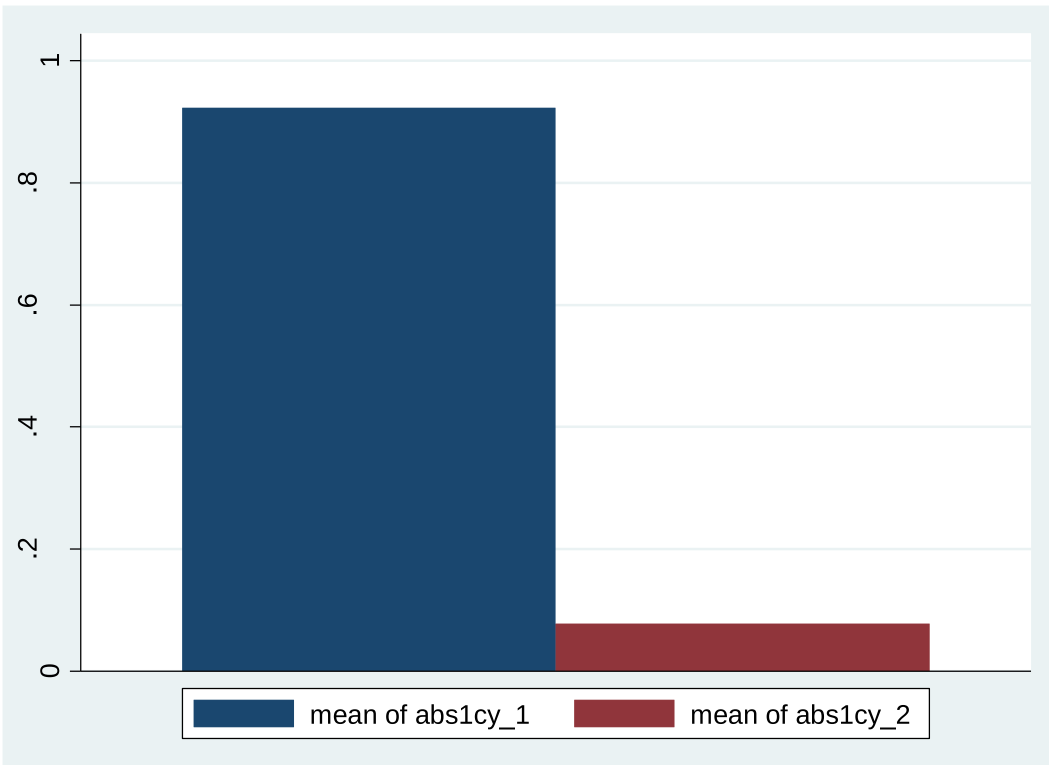


Variable name: *ABS1CY*
Source: Performance data set
Values: 0 or 1
Coding: a dummy variable equal to 1 if a player did not have performance data for the year prior to contract year and 0 if they did

The distribution of contracts by absentia one year prior to contract year is shown:

abs1cy	Freq.	Percent
0	322	92.26
1	27	7.74
Total	349	100.00

Frequency distribution of contracts by absentia one year prior to contract year in bar graph form:

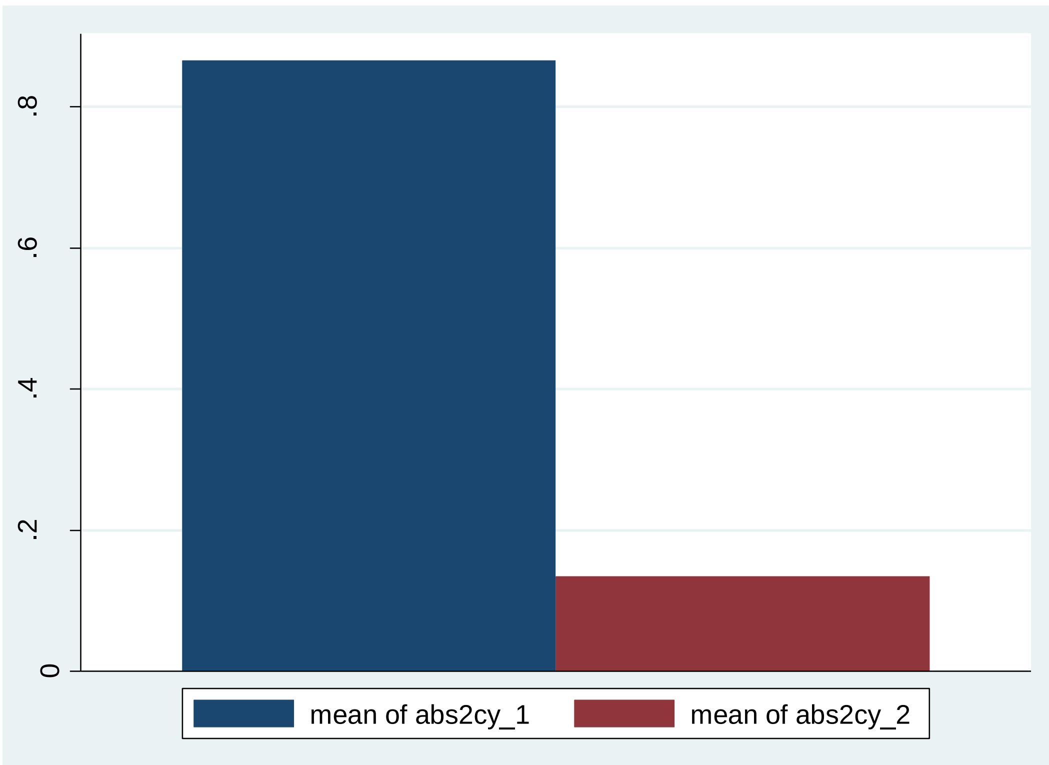


Variable name: *ABS2CY*
Source: Performance data set
Values: 0 or 1
Coding: a dummy variable equal to 1 if a player did not have performance data for the season two years prior to contract year and 0 if they did

The distribution of contracts by absentia two years prior to contract year is shown:

abs2cy	Freq.	Percent
0	302	86.53
1	47	13.47
Total	349	100.00

Frequency distribution of contracts by absentia two years prior to contract year in bar graph form:

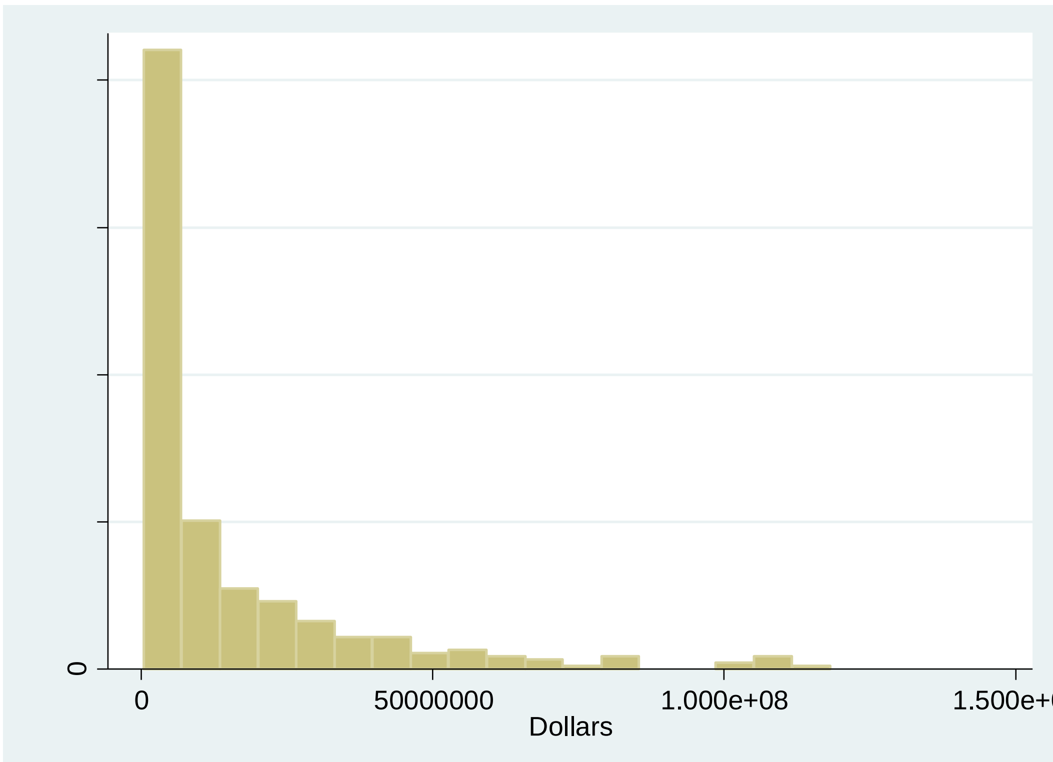


Variable name: *DOLLARS*
Source: Contract data set
Values: Integers ranging from 414,378 to 118,000,000
Coding: Total contract value reported in dollars

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
dollars	1.49e+07	2.18e+07	1352181	5500000	2.00e+07	414378	1.18e+08

Histogram:



Variable name: *lndollars*

Source: Contract data set

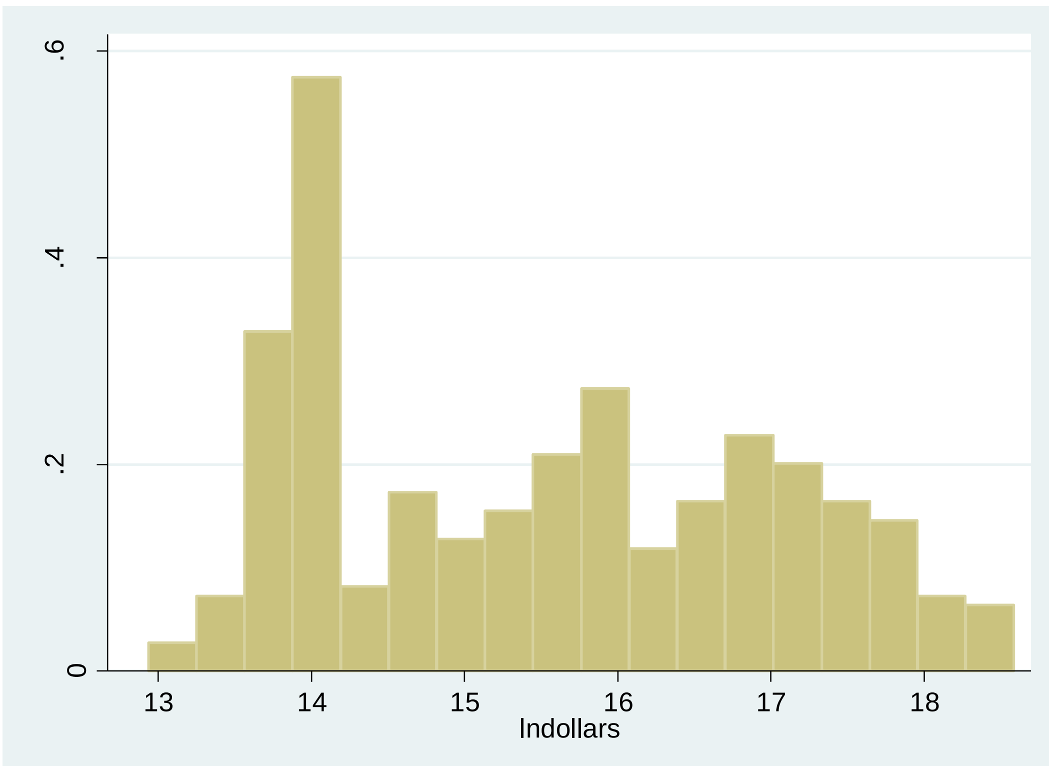
Values: Numbers ranging from 12.93 to 18.59

Coding: natural log of the total contract value reported in dollars

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lndollars	15.53283	1.461177	14.11723	15.52026	16.81124	12.93453	18.58789

Histogram:

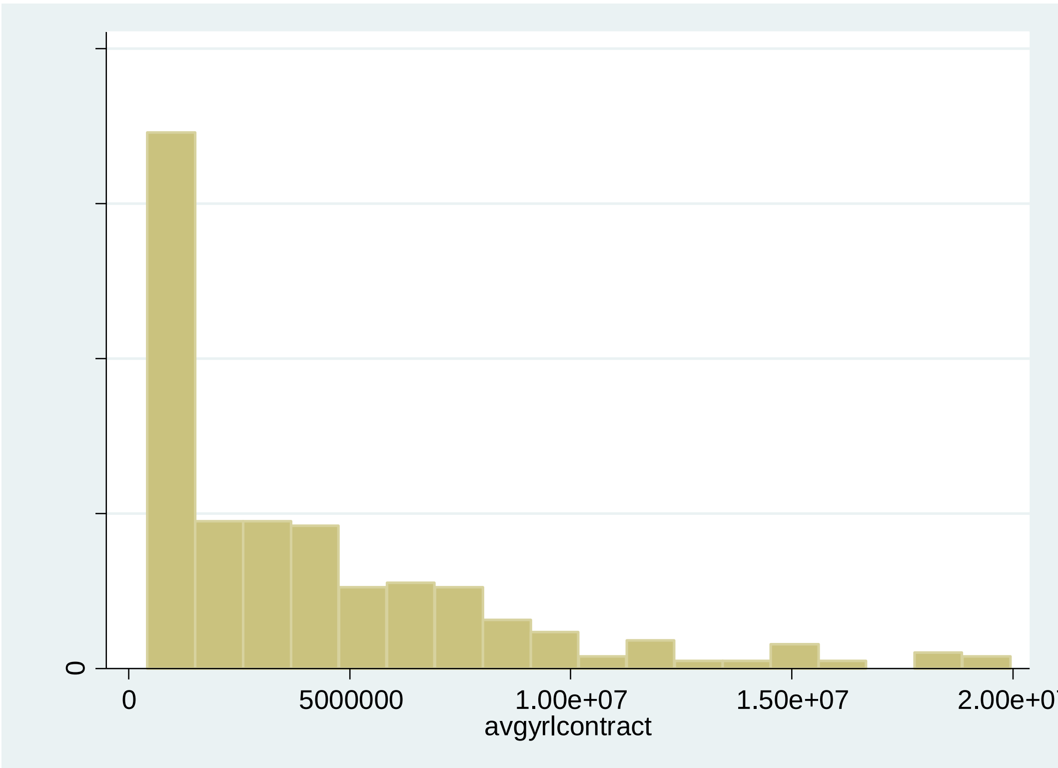


Variable name: *AvgYrlContract*
Source: Contract data set
Values: Numbers ranging from 414,378 to 19,900,000
Coding: Total contract value reported in dollars divided by the length of contract

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
avgyrlcont~t	4239869	4082067	1229255	2980000	6037500	414378	1.99e+07

Histogram:

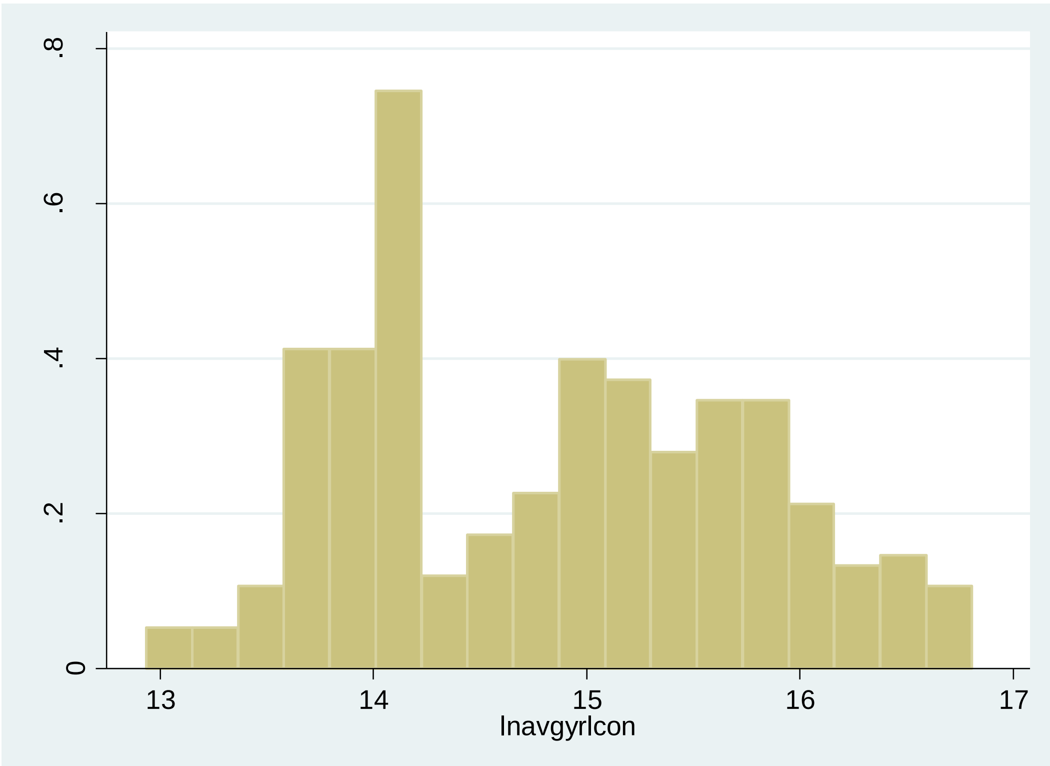


Variable name: *LnAvgYrlCon*
Source: Contract data set
Values: Numbers ranging from 12.93 to 16.81
Coding: Natural log of the total contract value reported in dollars divided by the length of contract

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnavgyrlcon	14.83859	.9278775	14.02192	14.90743	15.6135	12.93453	16.80868

Histogram:

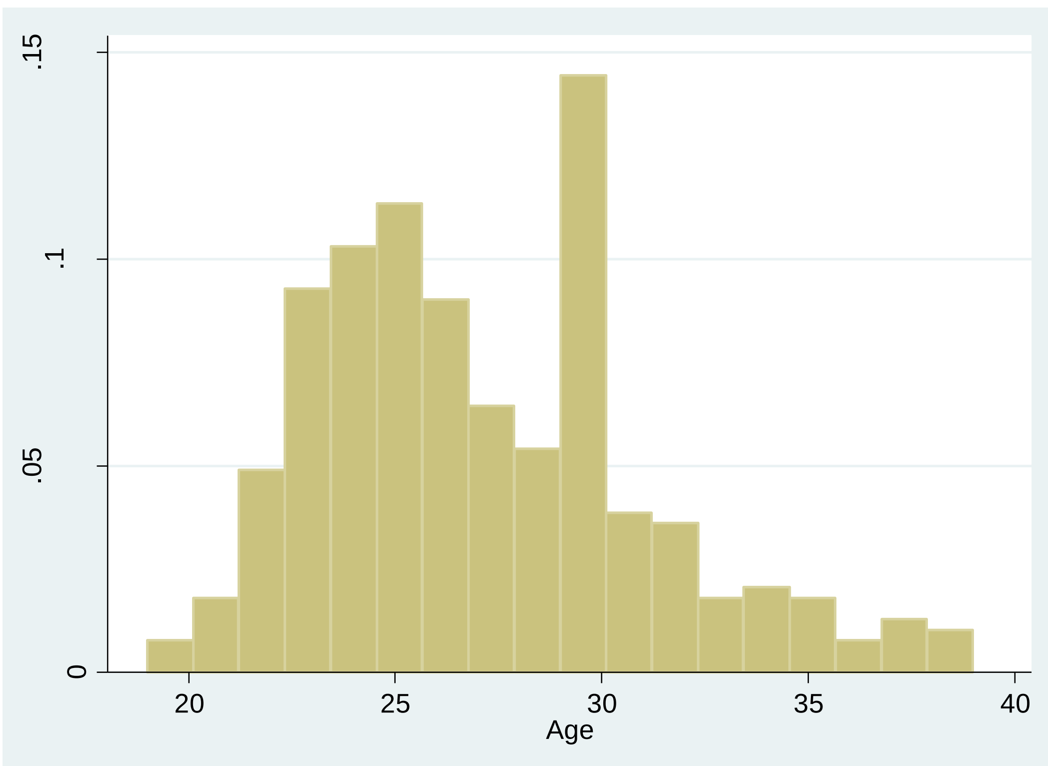


Variable name: *AICY*
Source: Contract data set
Values: Integers ranging from 19 to 39
Coding: age of a player in the year that they signed their free agent contract

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
aicy	27.04585	3.912574	24	26	29	19	39

Histogram:

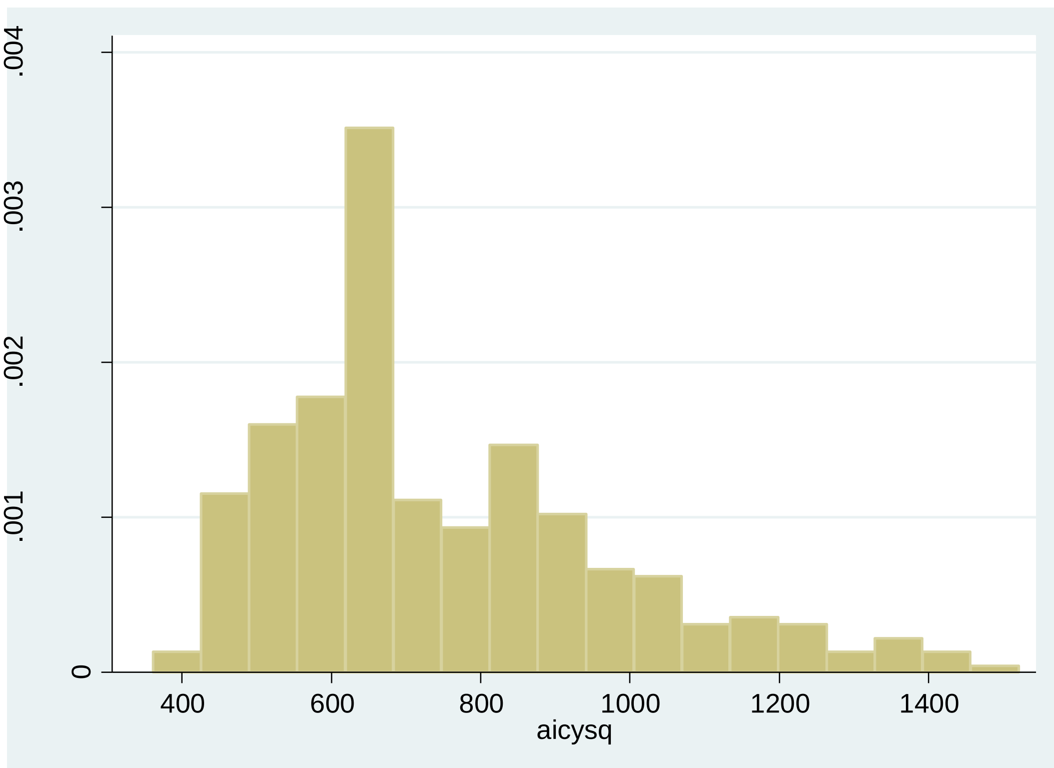


Variable name: *AICYSQ*
Source: Contract data set
Values: Integers ranging from to
Coding: age of a player in the year that they signed their free agent contract squared

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
aicysq	746.7421	223.2904	576	676	841	361	1521

Histogram:

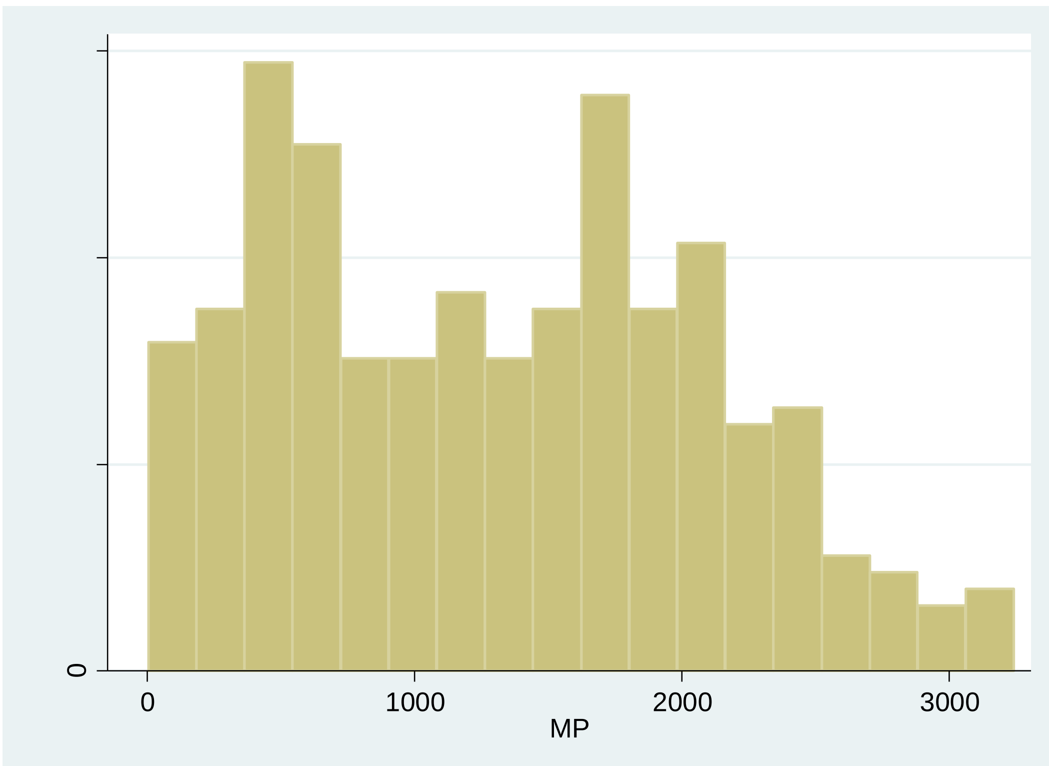


Variable name: *MPCY*
Source: Performance data set
Values: Integers ranging from 3 to 3,242
Coding: How many minutes of playing time a player logged during his contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mpcy	1305.946	794.4029	586	1295	1933	3	3242

Histogram:

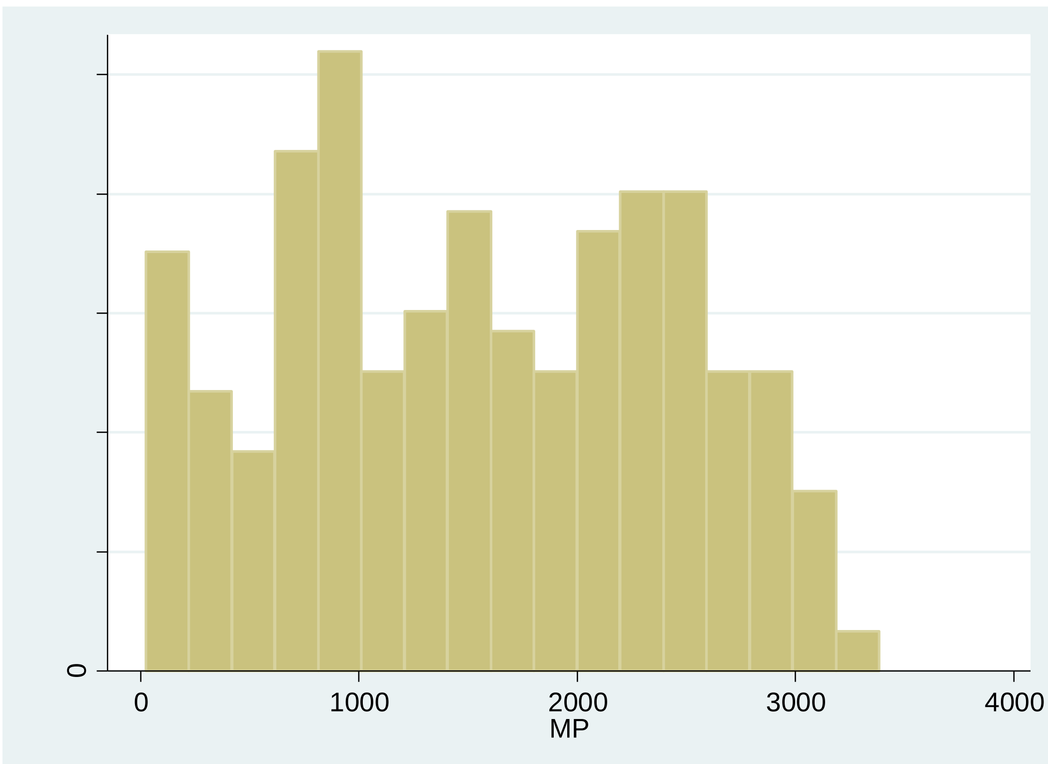


Variable name: *MP2CY*
Source: Performance data set
Values: Integers ranging from 23 to 3,384
Coding: How many minutes of playing time a player logged two years prior to contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mp2cy	1553.166	866.9173	858	1510	2276	23	3384

Histogram:

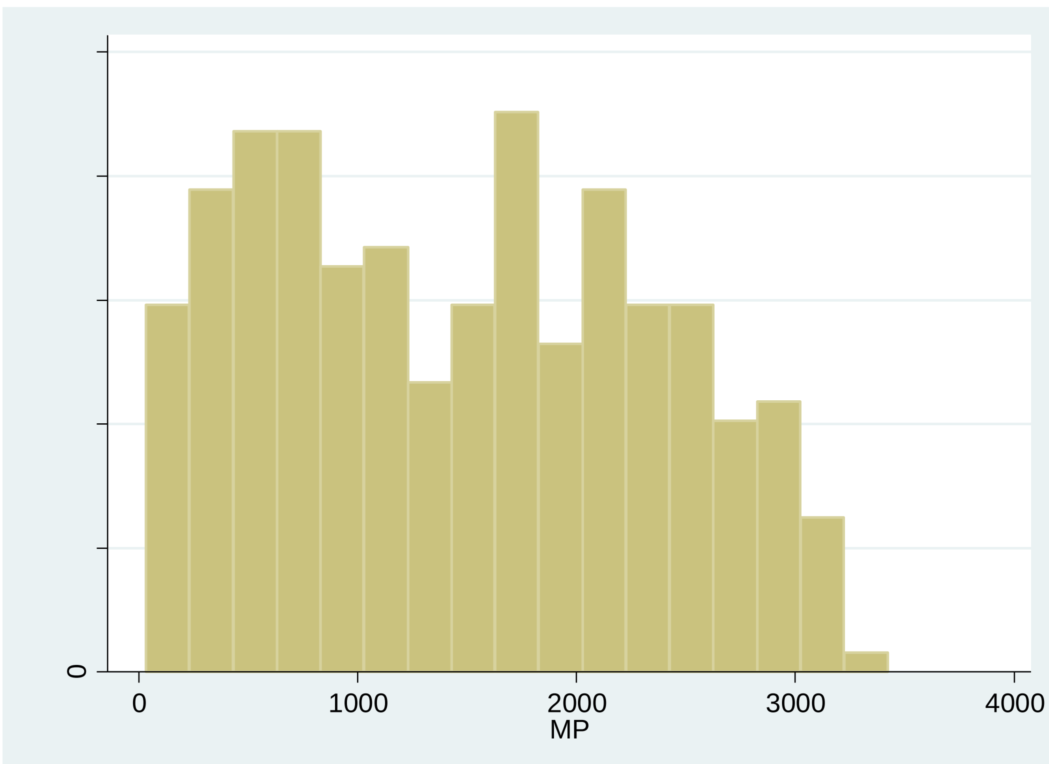


Variable name: *MP1CY*
Source: Performance data set
Values: Integers ranging from 34 to 3,424
Coding: How many minutes of playing time a player logged one year prior to contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mp1cy	1466.693	864.8361	685	1491	2193	34	3424

Histogram:

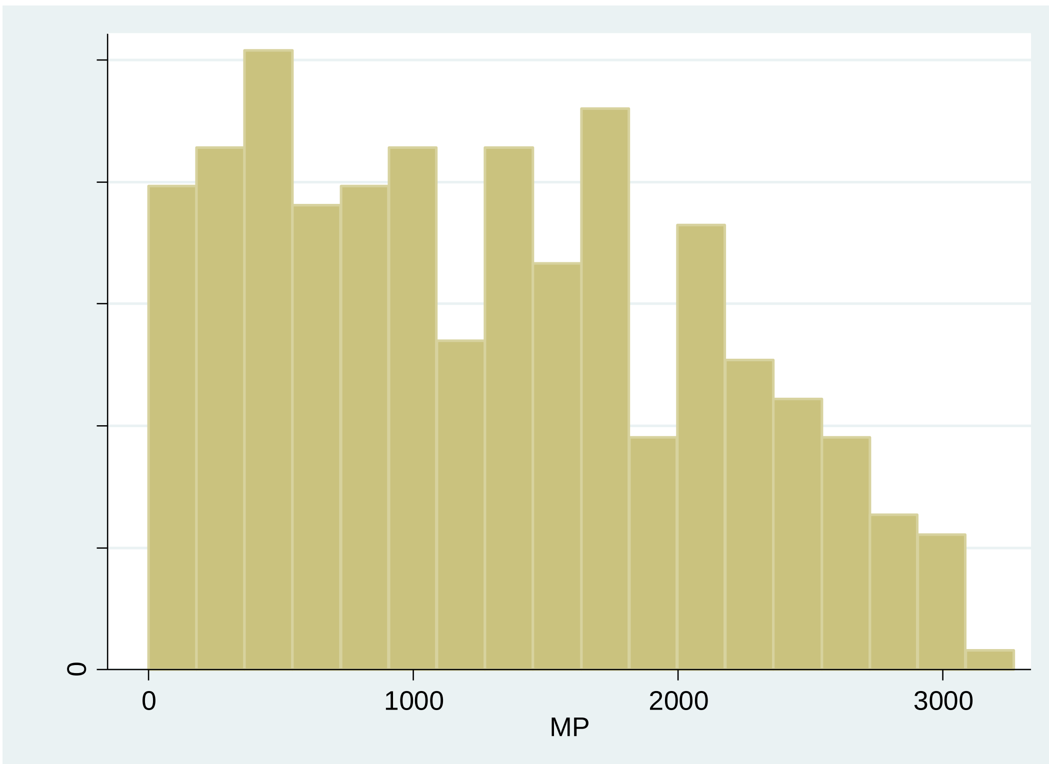


Variable name: *MPCY1*
Source: Performance data set
Values: Integers ranging from 0 to 3,269
Coding: How many minutes of playing time a player logged one year after contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mpcy1	1284.037	813.1768	575	1219	1892	0	3269

Histogram:

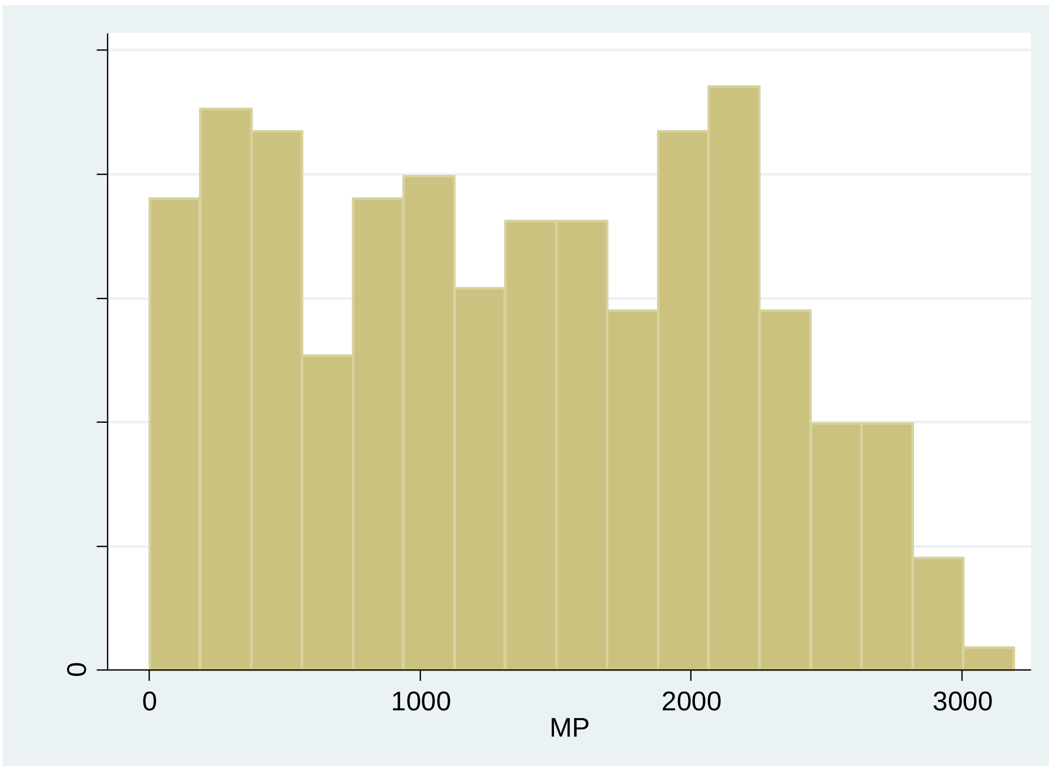


Variable name: *MPCY2*
Source: Performance data set
Values: Integers ranging from 1 to 3,193
Coding: How many minutes of playing time a player logged two years after contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mpcy2	1342.439	811.841	593	1354.5	2017	1	3193

Histogram:

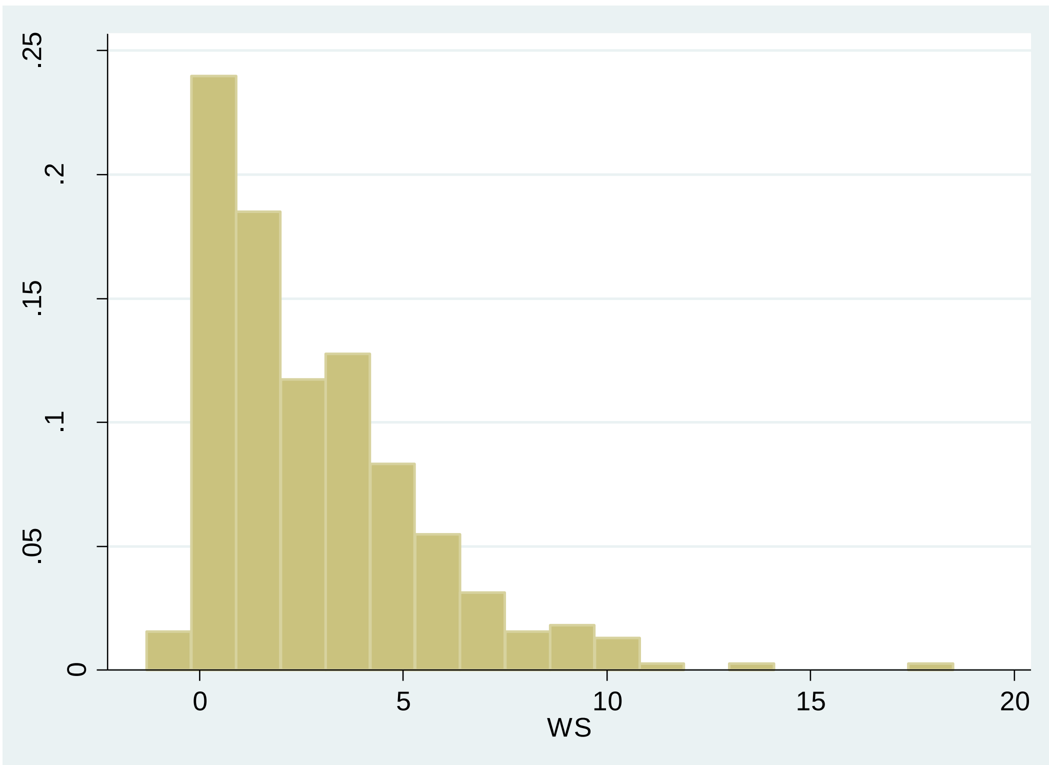


Variable name: *WSCY*
Source: Performance data set
Values: Numbers ranging from -1.3 to 18.5
Coding: The win share value calculated for each player during their contract year season

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy	2.805444	2.686396	.7	2.2	4.1	-1.3	18.5

Histogram:

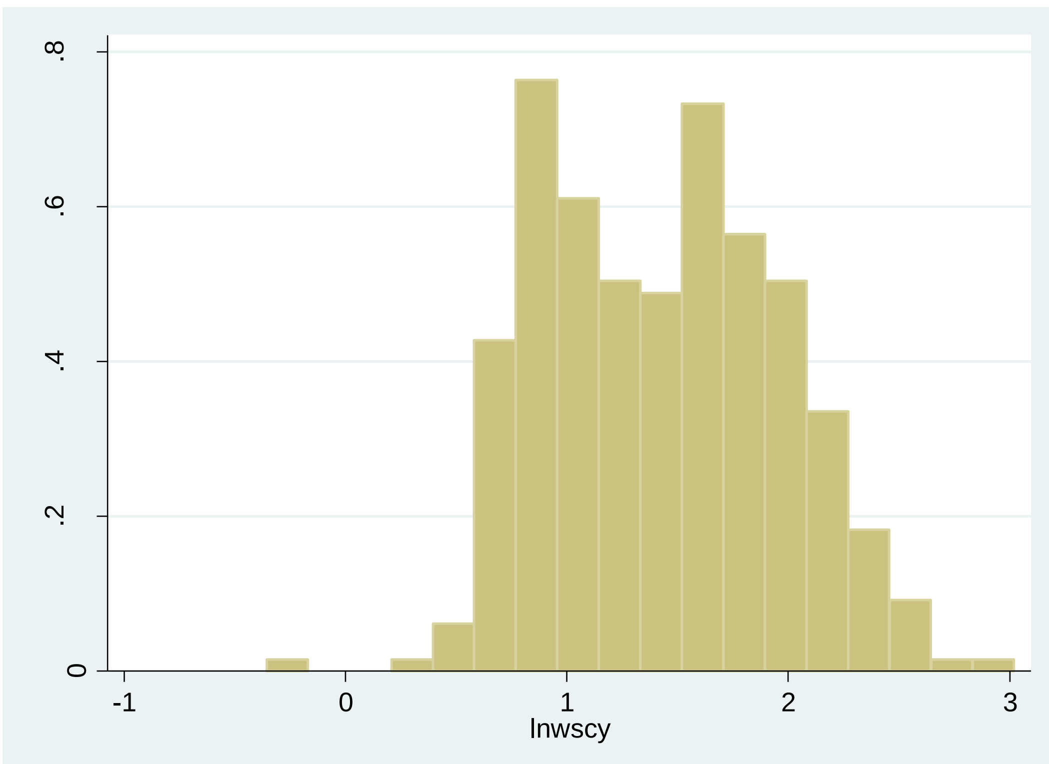


Variable name: *LNWSCY*
Source: Performance data set
Values: Numbers ranging from -0.35 to 3.02
Coding: The natural log of the win share value calculated for each player during their contract year season plus 2 (to keep all numbers feasible for natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy	1.429839	.5279245	.9932518	1.435085	1.808289	-.3566749	3.020425

Histogram:

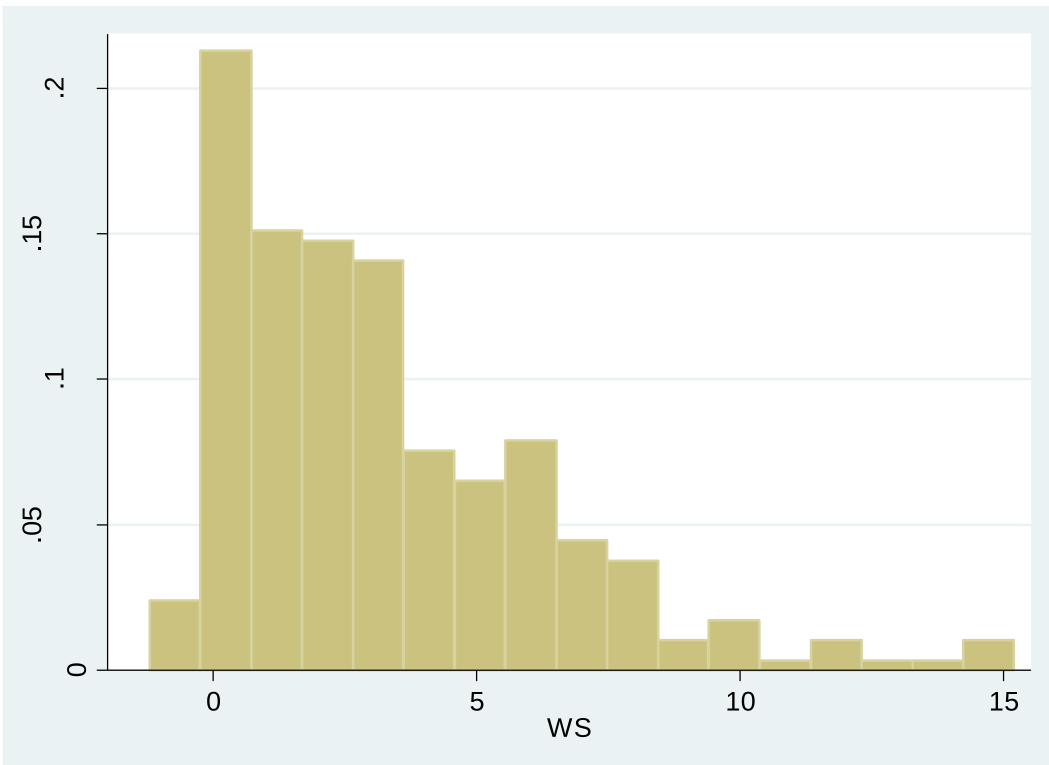


Variable name: *WS2CY*
Source: Performance data set
Values: Numbers ranging from -1.2 to 15.2
Coding: The win share value calculated for each player during the season two years prior to their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
ws2cy	3.24404	3.043453	.9	2.45	5	-1.2	15.2

Histogram:



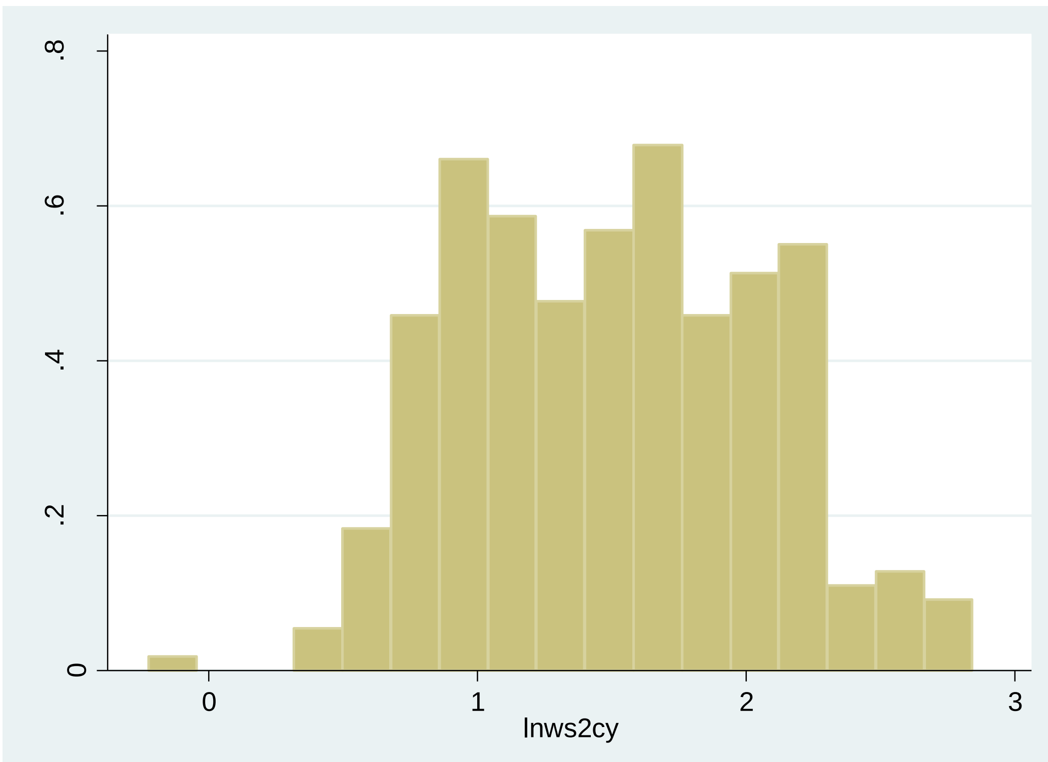
Variable name: *WS2CYA*
Source: Performance data set
Values: Numbers ranging from -1.2 to 15.2
Coding: the win share value calculated for each player during the season two years prior to contract year with zeroes filled in for the missing values (this is done for modeling purposes)

Variable name: *LNWS2CY*
Source: Performance data set
Values: Numbers ranging from -.22 to 2.85
Coding: the natural log of the win share value calculated for each player during the season two years prior to their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnws2cy	1.502765	.5567775	1.064711	1.492841	1.94591	-.2231435	2.844909

Histogram:



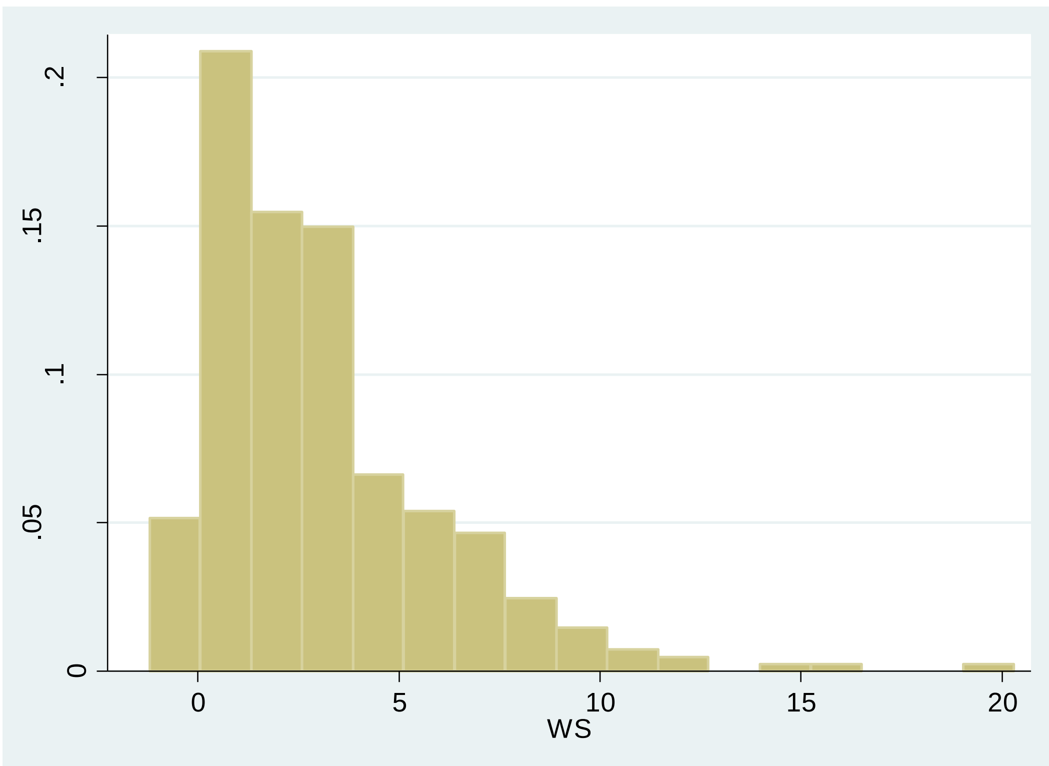
Variable name: *LnWS2CYA*
Source: Performance data set
Values: Numbers ranging from -1.2 to 15.2
Coding: the natural log of the win share value calculated for each player during the season two years prior to contract year with zeroes filled in for the missing values (this is done for modeling purposes)

Variable name: *WS1CY*
Source: Performance data set
Values: Numbers ranging from -1.2 to 20.3
Coding: The win share value calculated for each player during the season one year prior to their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
ws1cy	3.051242	2.942843	.9	2.4	4.4	-1.2	20.3

Histogram:



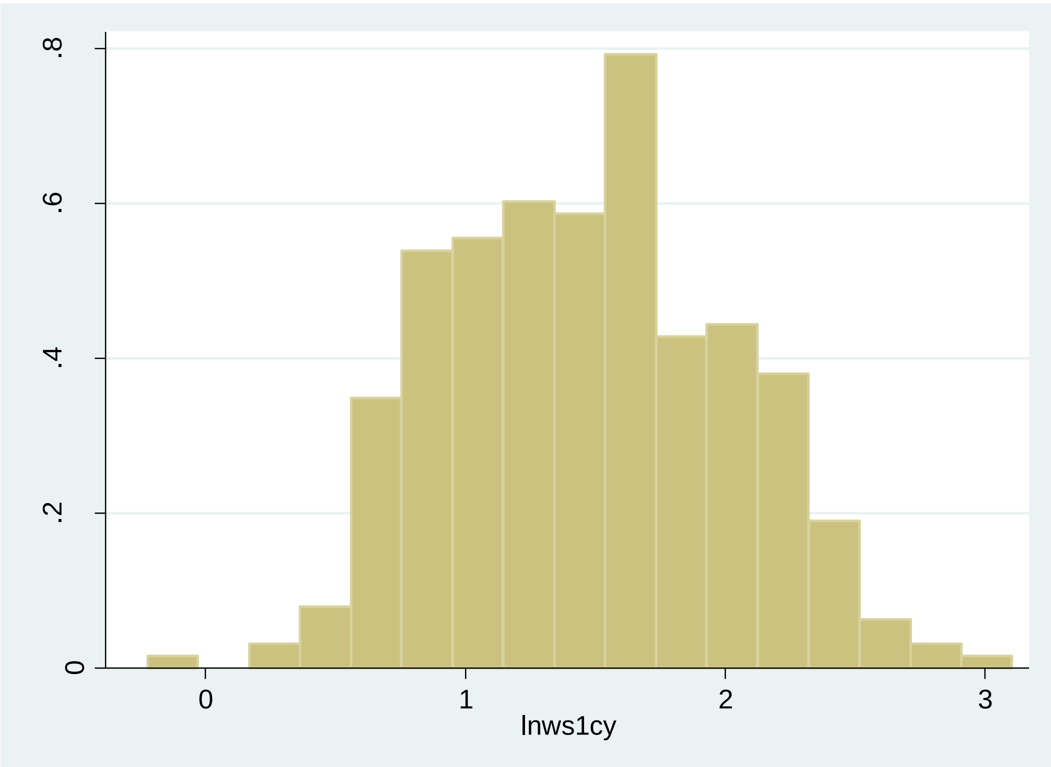
Variable name: *WS1CYA*
Source: Performance data set
Values: Numbers ranging from -1.2 to 20.3
Coding: the win share value calculated for each player during the season one year prior to contract year with zeroes filled in for the missing values (this is done for modeling purposes)

Variable name: *LNWS1CY*
Source: Performance data set
Values: Numbers ranging from -0.22 to 3.11
Coding: the natural log of the win share value calculated for each player during the season one year prior to their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnws1cy	1.471514	.5425029	1.064711	1.481605	1.856298	-.2231435	3.104587

Histogram:



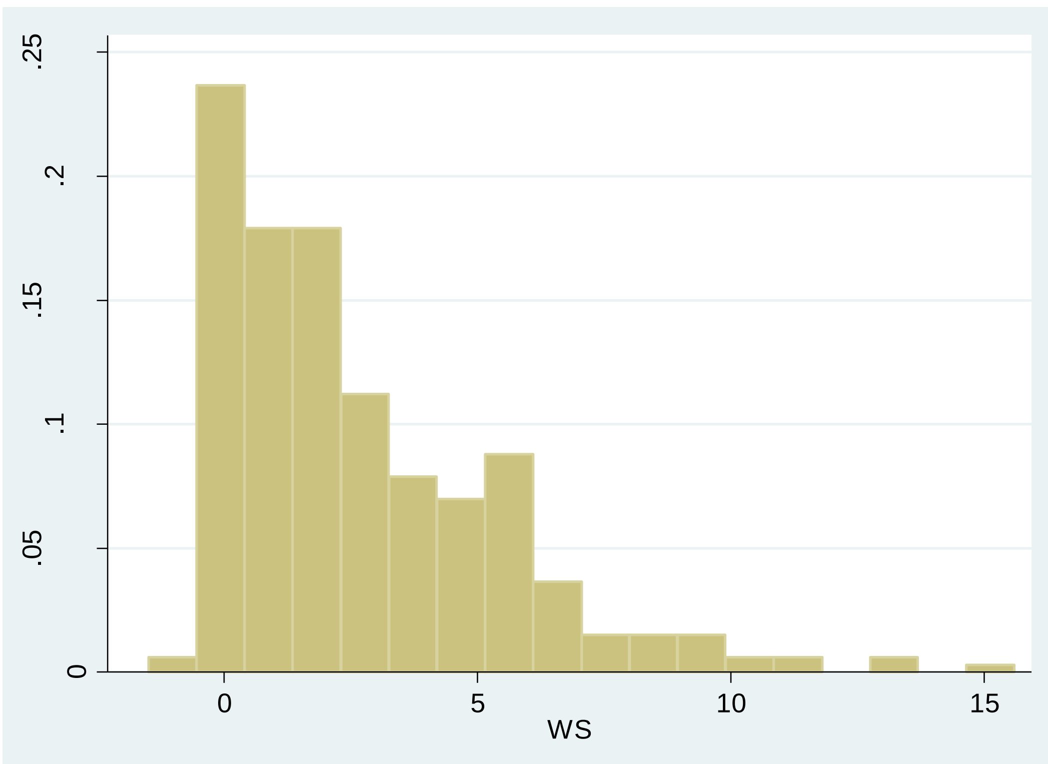
Variable name: *LnWS1CYA*
Source: Performance data set
Values: Numbers ranging from -1.2 to 15.2
Coding: the natural log of the win share value calculated for each player during the season one year prior to contract year with zeroes filled in for the missing values (this is done for modeling purposes)

Variable name: *WSCY1*
Source: Performance data set
Values: Numbers ranging from -1.5 to 15.6
Coding: the win share value calculated for each player during the season one year after their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy1	2.669164	2.67144	.5	1.9	4.1	-1.5	15.6

Histogram:

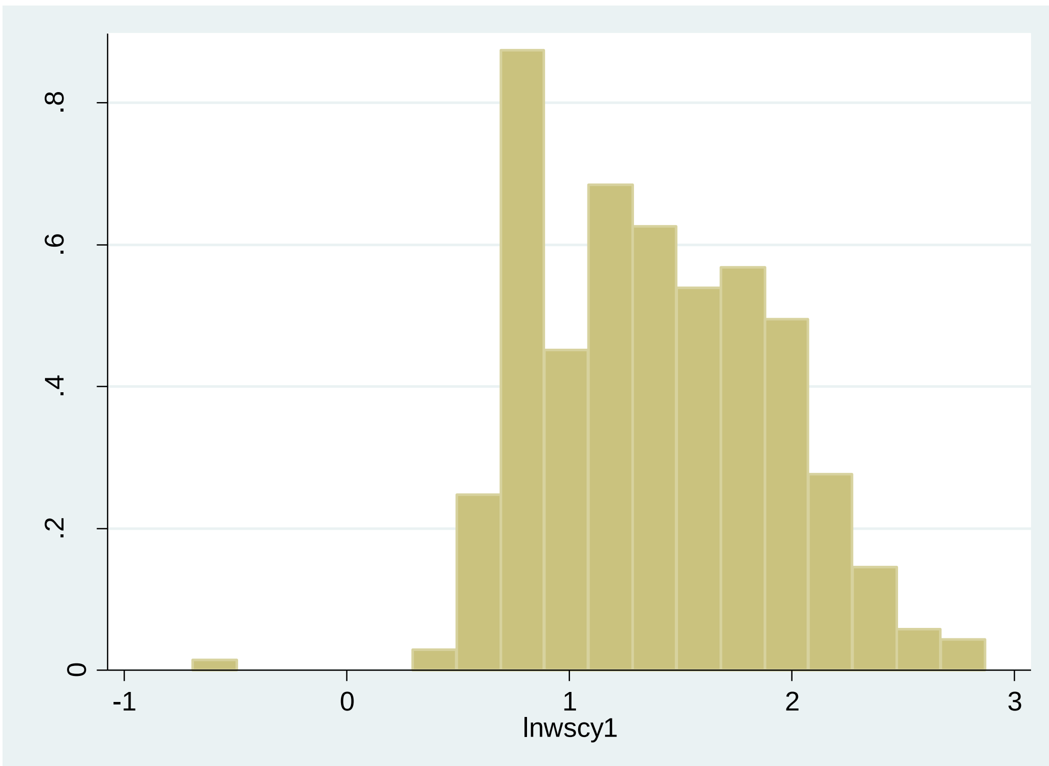


Variable name: *LNWSCY1*
Source: Performance data set
Values: Numbers ranging from -0.69 to 2.87
Coding: the natural log of the win share value calculated for each player during the season one year after their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy1	1.39574	.5365741	.9162908	1.360977	1.808289	-.6931472	2.867899

Histogram:

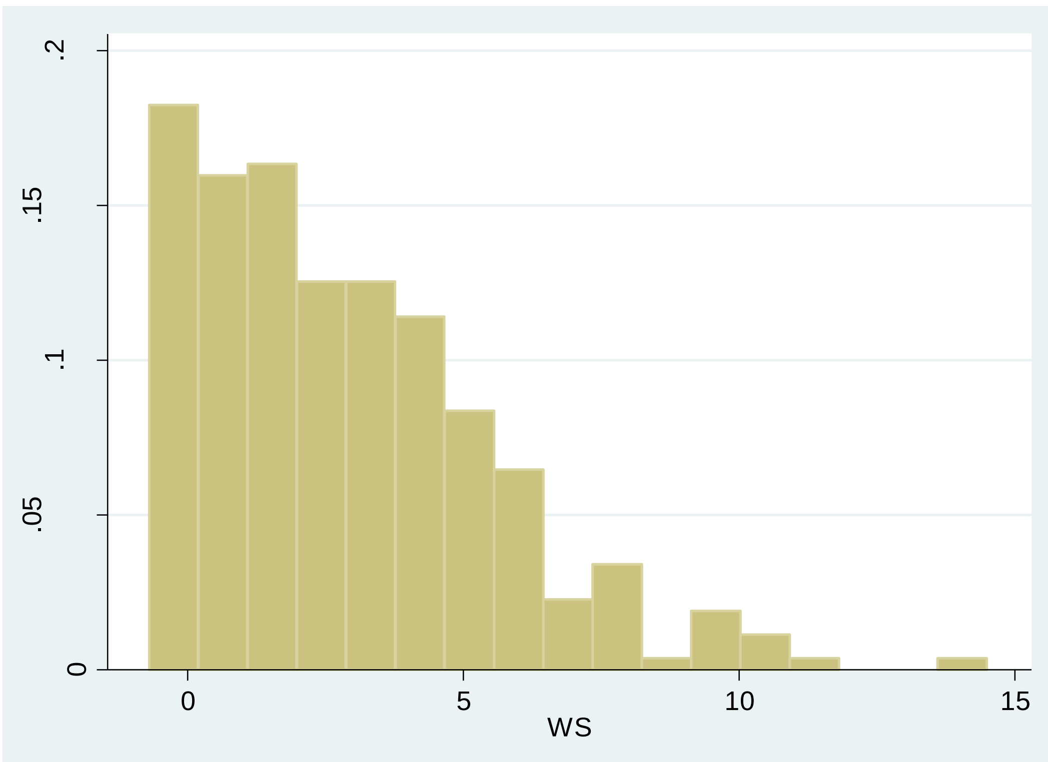


Variable name: *WSCY2*
Source: Performance data set
Values: Numbers ranging from -.7 to 14.5
Coding: the win share value calculated for each player during the season two years after their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy2	2.852721	2.620393	.6	2.3	4.4	-.7	14.5

Histogram:

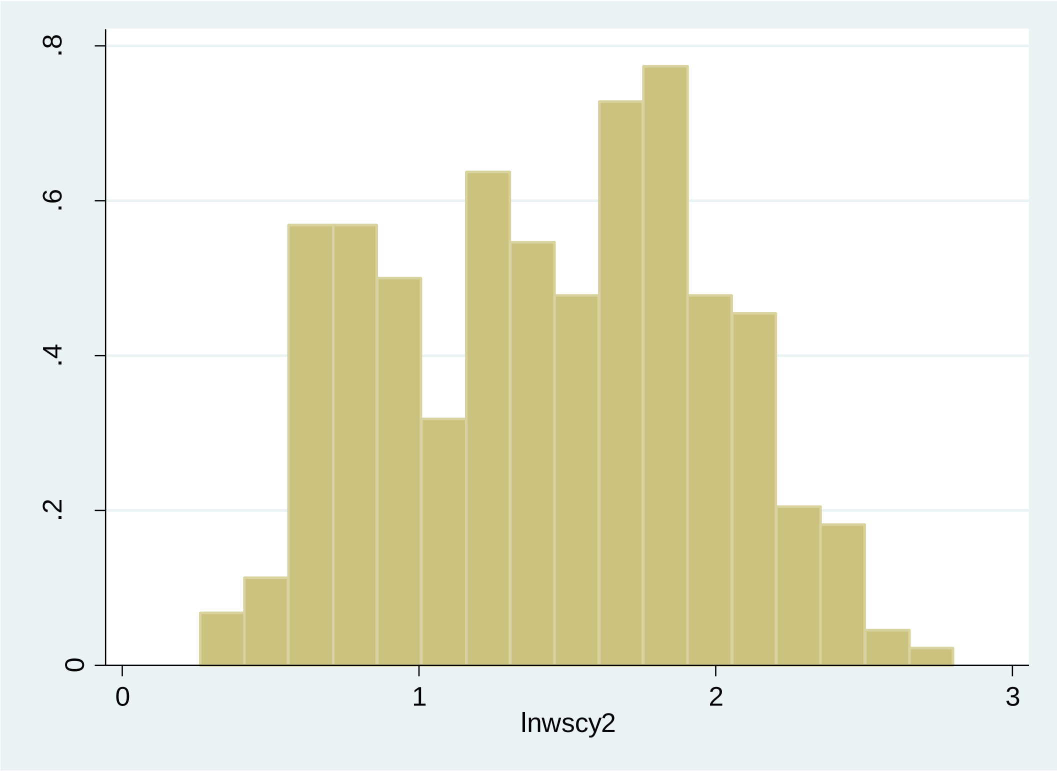


Variable name: *LNWSCY2*
Source: Performance data set
Values: Numbers ranging from .26 to 2.81
Coding: the natural log of the win share value calculated for each player during the season two years after their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy2	1.438026	.5383272	.9555115	1.458615	1.856298	.2623643	2.80336

Histogram:

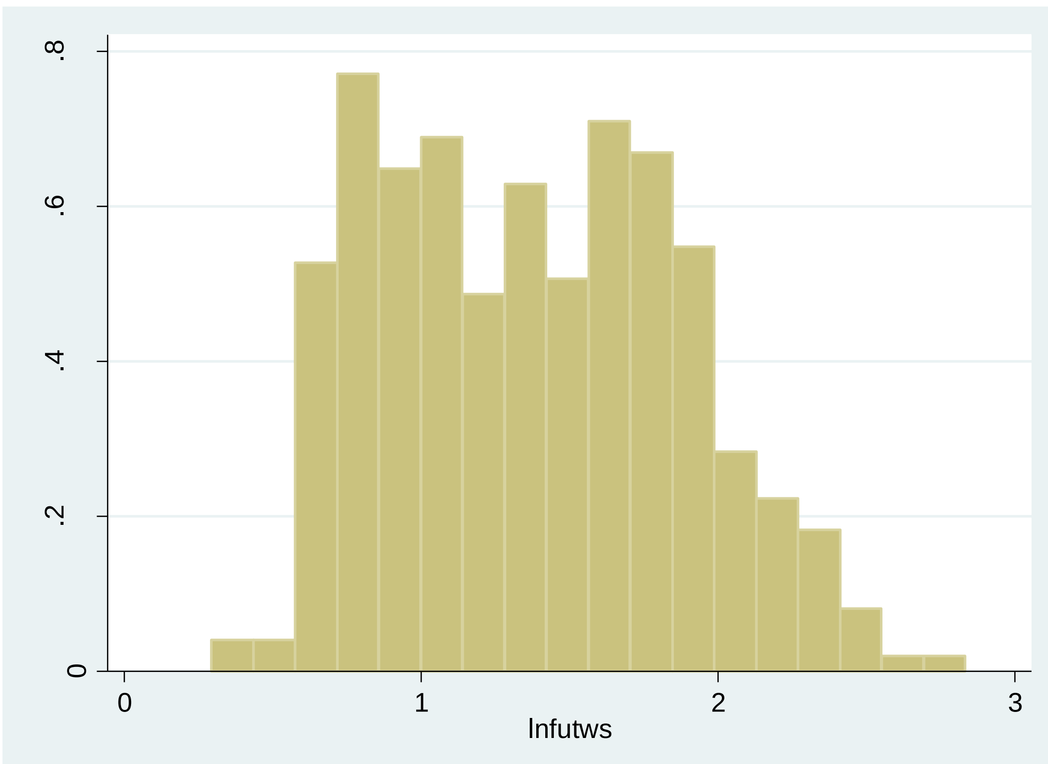


Variable name: *LnFutWS*
Source: Performance data set
Values: Numbers ranging from .29 to 2.84
Coding: The average of lnwscy1 and lnwscy2 or one or the other if both are not observed in the data

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnfutws	1.37579	.4987678	.9488099	1.32943	1.762002	.2938933	2.83563

Histogram:



Methodology B:

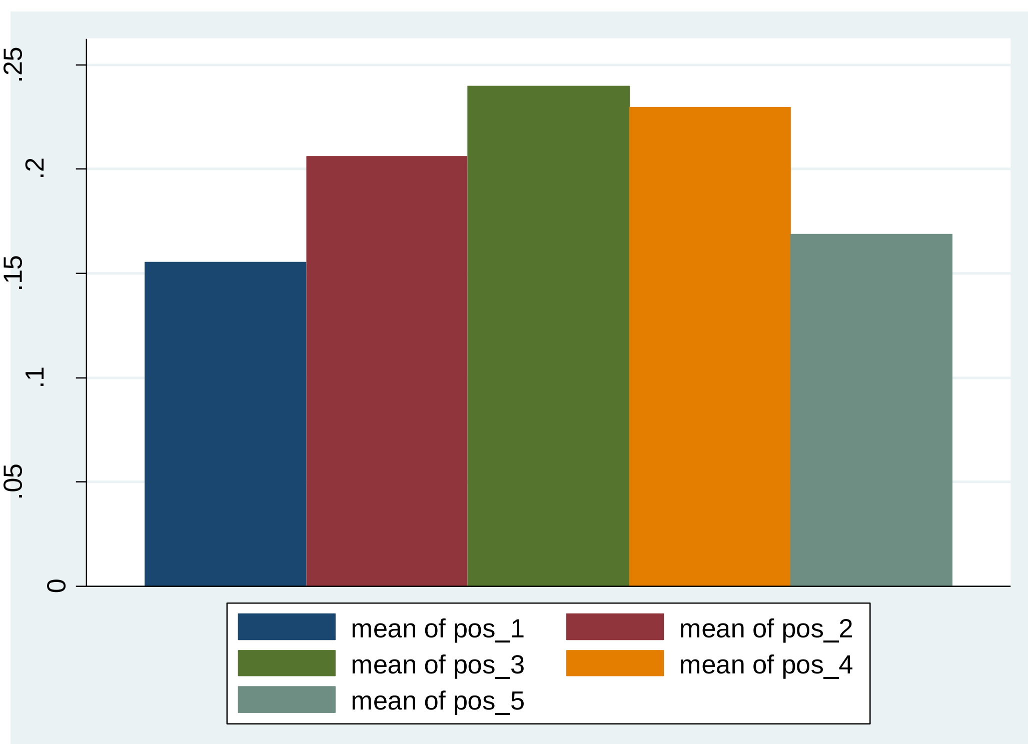
This data set includes National Basketball Association contracts signed by free agents between the 2006-2007 and 2011-2012 seasons. There are 296 NBA contracts in this data set which contains information relating their performance across the three years prior and the duration of their contract up to six years later. There are 207 unique players with 62 players appearing more than one time.

Variable name: *POS*
Source: Contract data set
Values: each of the five positions- Center, Power Forward, Point Guard, Small Forward, Shooting Guard

The distribution of contracts across positions is shown.

Pos	Freq.	Percent
C	46	15.54
PF	61	20.61
PG	71	23.99
SF	68	22.97
SG	50	16.89
Total	296	100.00

Frequency distribution of each position in bar graph form:

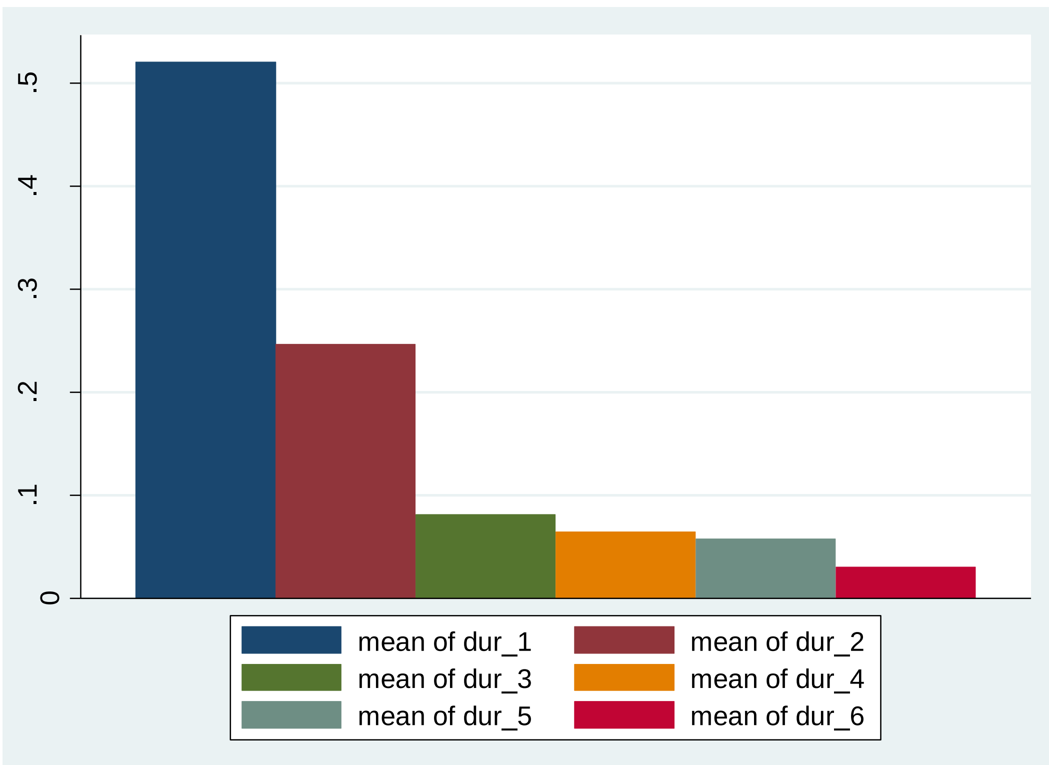


Variable name: *DURATION*
Source: Contract data set
Values: integers from 1 through 6
Coding: Length of contract in years

The distribution of contract durations is shown.

duration	Freq.	Percent	Cum.
1	154	52.03	52.03
2	73	24.66	76.69
3	24	8.11	84.80
4	19	6.42	91.22
5	17	5.74	96.96
6	9	3.04	100.00
Total	296	100.00	

Frequency distribution of each contract duration in bar graph form:

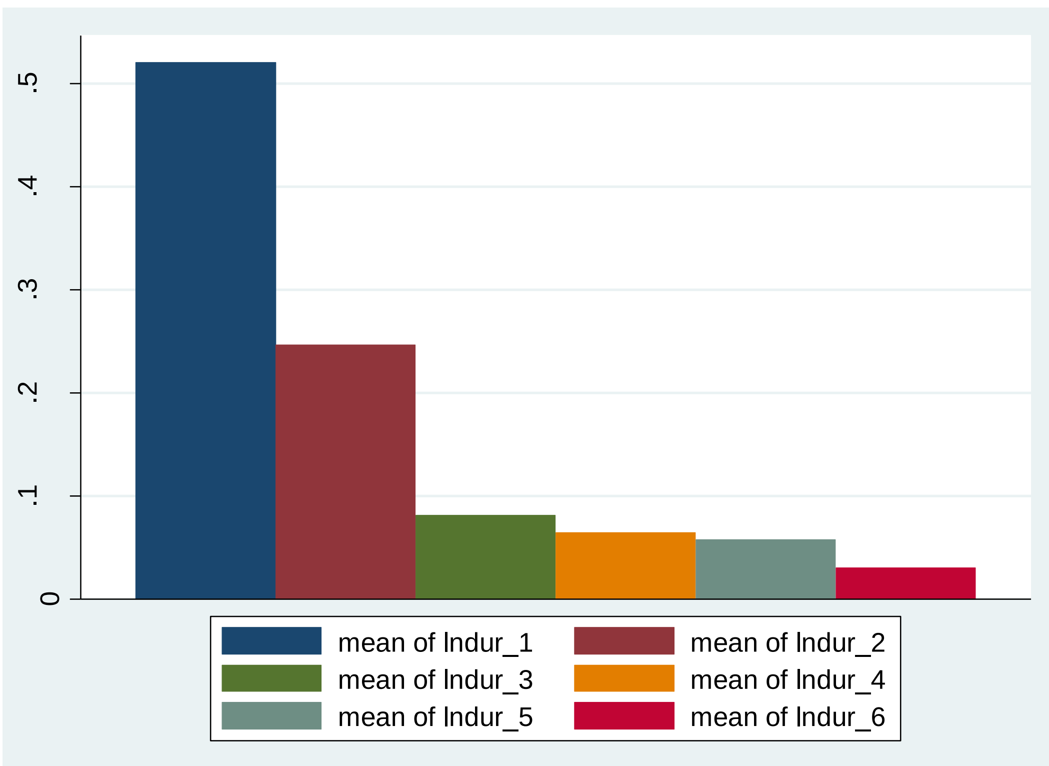


Variable name: *LnDURATION*
Source: Contract data set
Values: Numbers ranging from 0 to 1.8
Coding: Natural log of the length of contract in years

The distribution of the natural log of contract durations is shown.

lnduration	Freq.	Percent	Cum.
0	154	52.03	52.03
.6931472	73	24.66	76.69
1.098612	24	8.11	84.80
1.386294	19	6.42	91.22
1.609438	17	5.74	96.96
1.791759	9	3.04	100.00
Total	296	100.00	

Frequency distribution of the natural log of each contract duration in bar graph form:



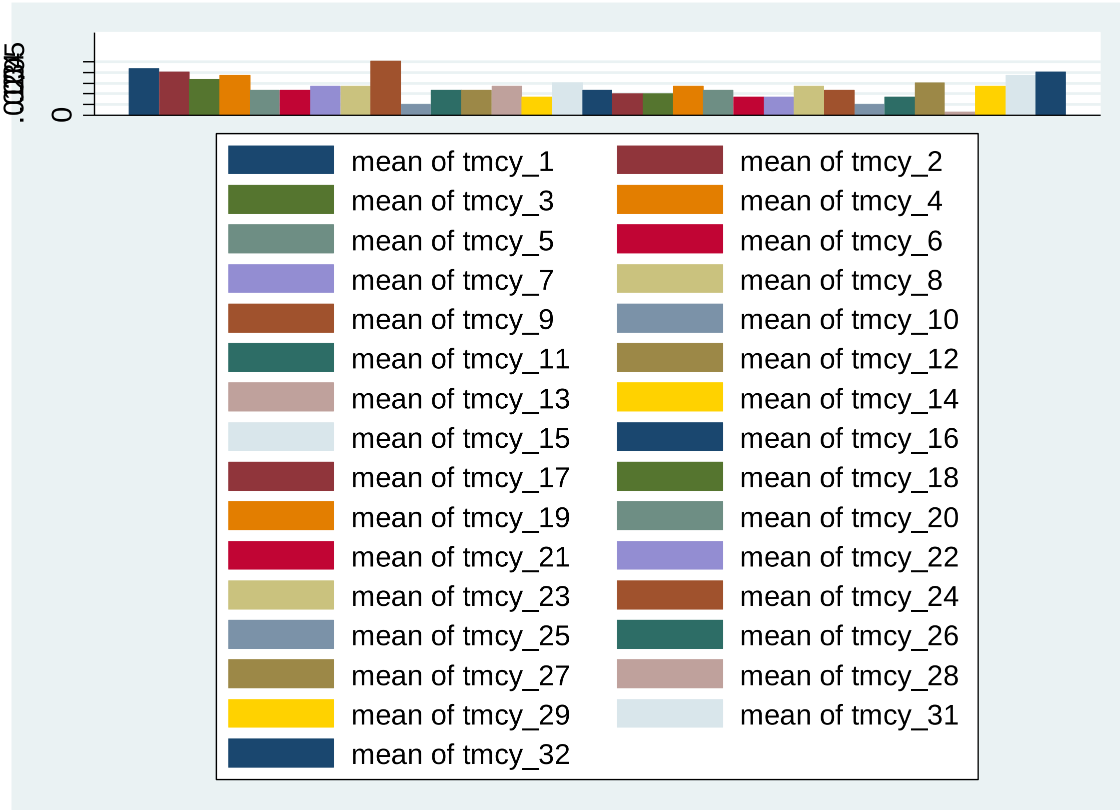
Variable name: *TMCY*
Source: Contract data set
Values: Each of the 31 different NBA teams over the period in question. *Players who switched teams during a season are not included here.

The distribution of contracts by team is shown:

Tm	Freq.	Percent
ATL	13	4.39
BOS	12	4.05
CHA	10	3.38
CHI	11	3.72
CLE	7	2.36
DAL	7	2.36
DEN	8	2.70
DET	8	2.70
GSW	15	5.07
HOU	3	1.01
IND	7	2.36
LAC	7	2.36
LAL	8	2.70
MEM	5	1.69
MIA	9	3.04
MIL	7	2.36
MIN	6	2.03
NJN	6	2.03
NOH	8	2.70
NYK	7	2.36
OKC	5	1.69
ORL	5	1.69
PHI	8	2.70
PHO	7	2.36
POR	3	1.01
SAC	5	1.69
SAS	9	3.04
SEA	1	0.34
TOR	8	2.70
UTA	11	3.72
WAS	12	4.05
Total	238	80.41

**Note that the team category for “TOT” was left out as it represents when a player was traded midseason and it is not an actual NBA team.

Frequency distribution of each team during contract year in bar graph form:



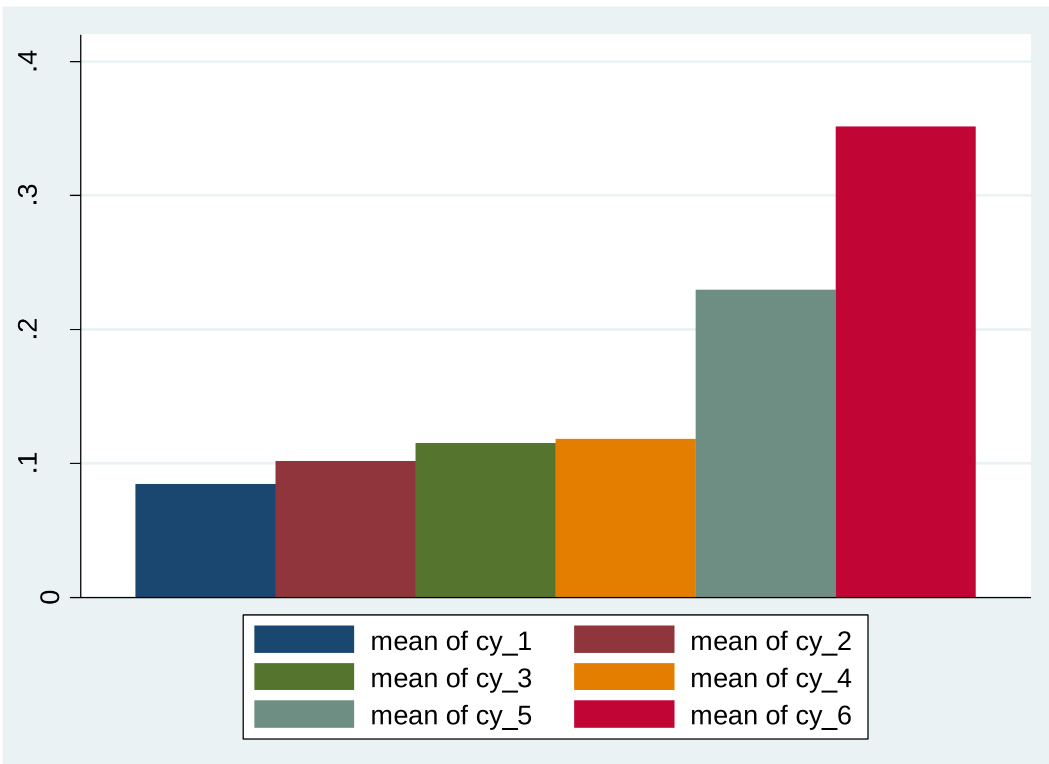
**Note that the team category for "TOT" was left out as it represents when a player was traded midseason and it is not an actual NBA team. This would be represented by tmcy_30.

Variable name: *CY*
Source: Contract data set
Values: Integers ranging from 2007 to 2012
Coding: the year that a player signed his contract

The distribution of contracts by contract year is shown:

cy	Freq.	Percent	Cum.
2007	25	8.45	8.45
2008	30	10.14	18.58
2009	34	11.49	30.07
2010	35	11.82	41.89
2011	68	22.97	64.86
2012	104	35.14	100.00
Total	296	100.00	

Frequency distribution of contract by contract year in bar graph form:

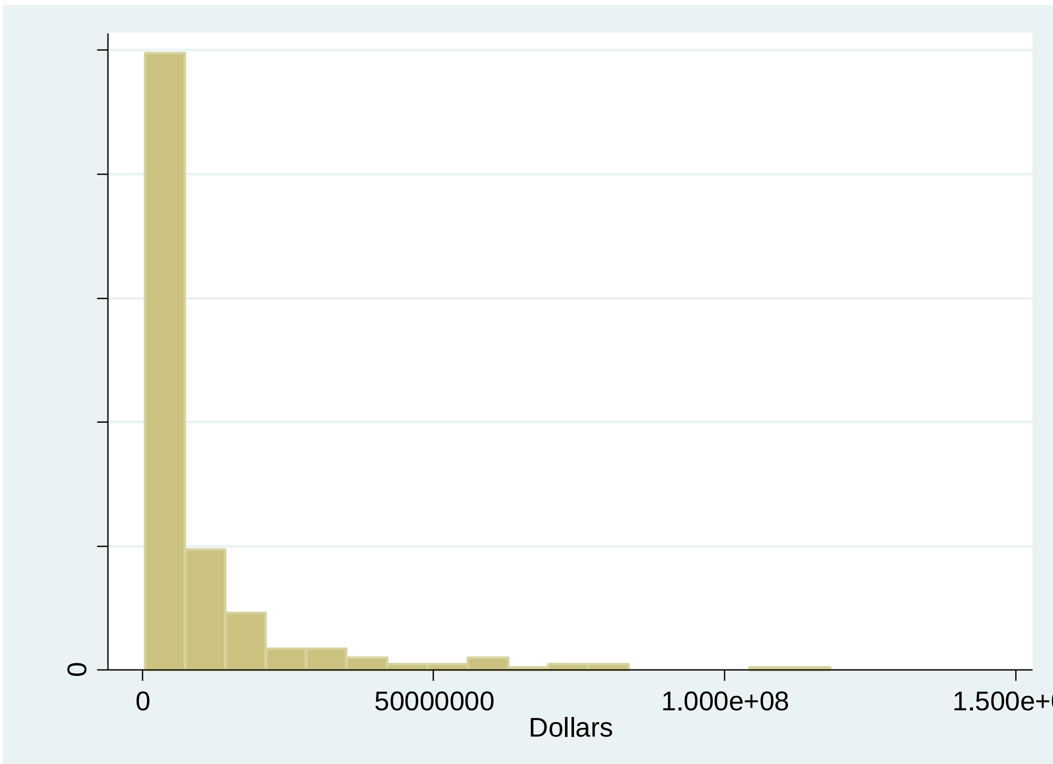


Variable name: *DOLLARS*
Source: Contract data set
Values: Integers ranging from 414,378 to 118,000,000
Coding: Total contract value reported in dollars

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
dollars	9338297	1.66e+07	1164070	2500000	9000000	414378	1.18e+08

Histogram:



Variable name: *LnDOLLARS*

Source: Contract data set

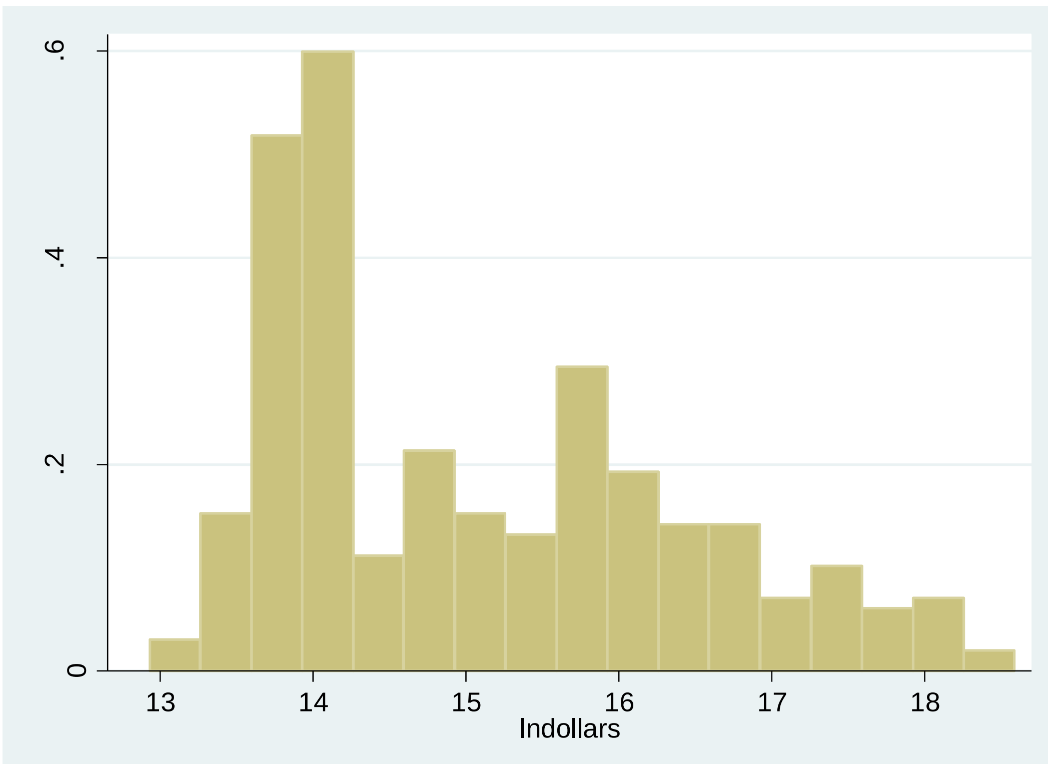
Values: Numbers ranging from 12.93 to 18.59

Coding: natural log of the total contract value reported in dollars

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lndollars	15.0728	1.326588	13.96732	14.7318	16.01274	12.93453	18.58789

Histogram:

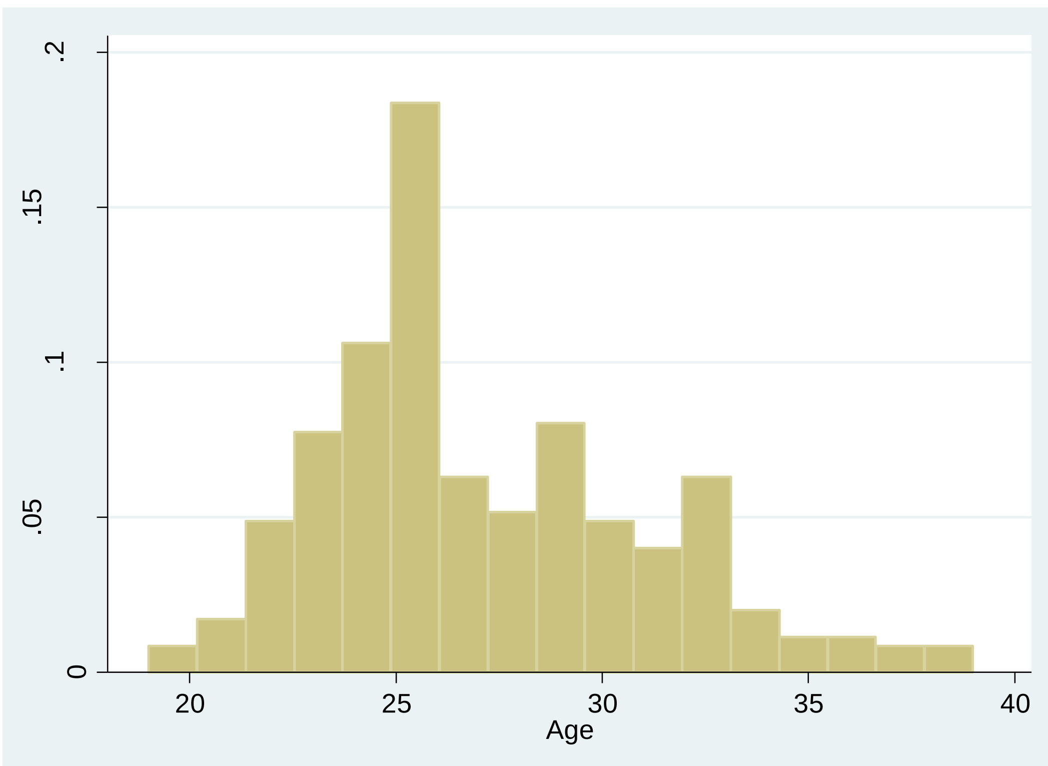


Variable name: *AICY*
Source: Contract data set
Values: Integers ranging from 19 to 39
Coding: Age of a player in the year that they signed their free agent contract

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
aicy	27.06081	3.902587	24	26	29.5	19	39

Histogram:

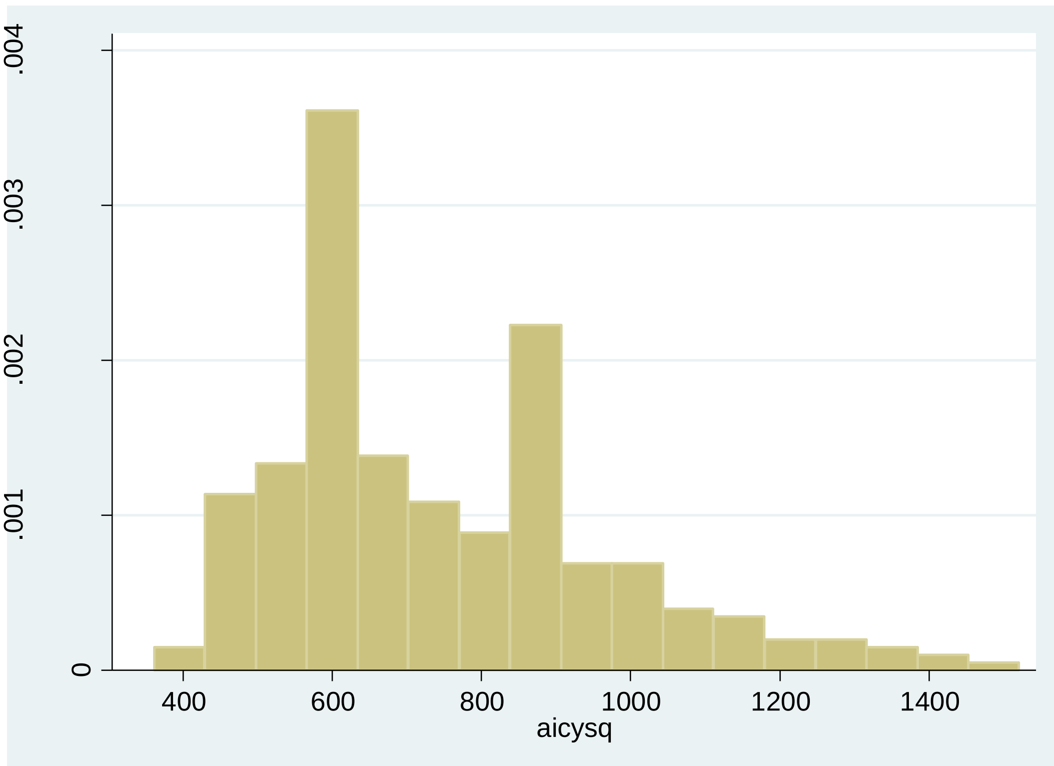


Variable name: *AICYSQ*
Source: Contract data set
Values: Integers ranging from 361 to 1521
Coding: Age of a player in the year that they signed their free agent contract squared

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
aicysq	747.4662	221.8529	576	676	870.5	361	1521

Histogram:

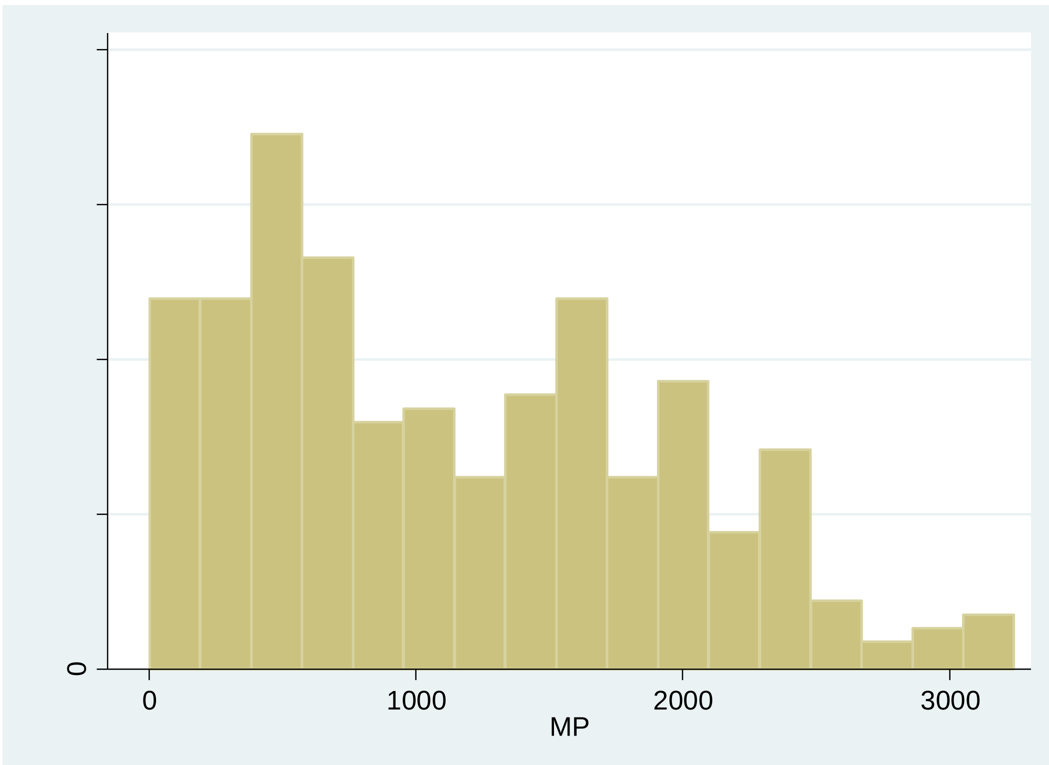


Variable name: *MPCY*
Source: Performance data set
Values: Integers ranging from 3 to 3,242
Coding: How many minutes of playing time a player logged during his contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mpcy	1155.257	783.2175	474.5	1031	1723.5	3	3242

Histogram:

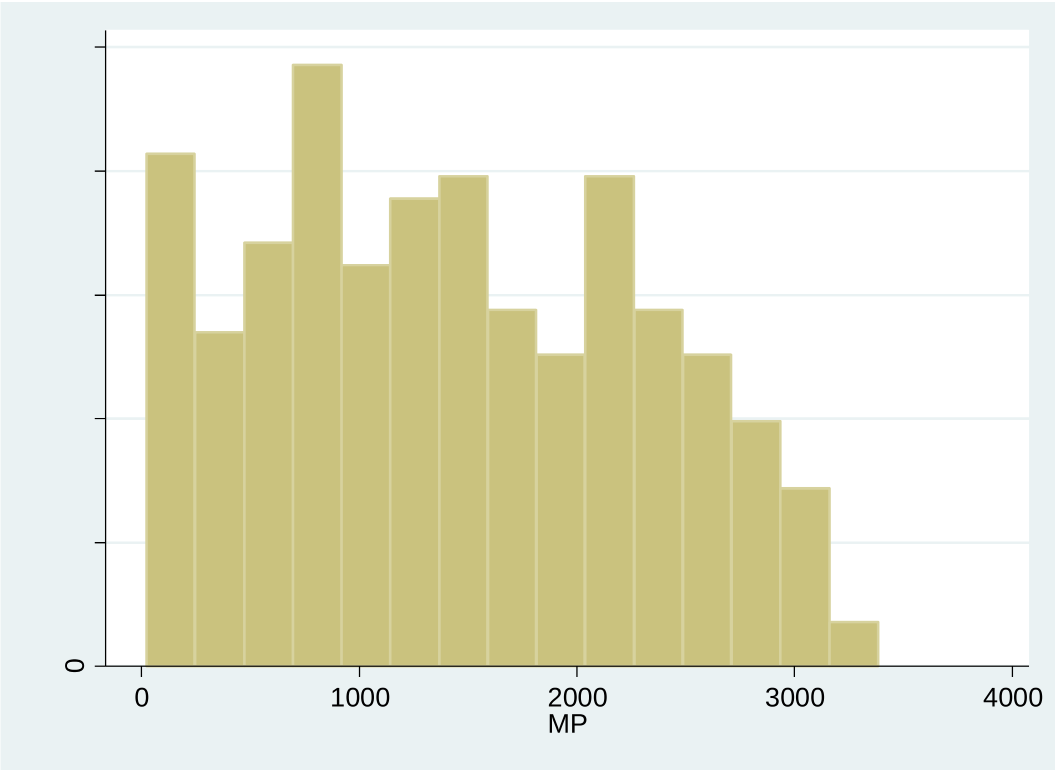


Variable name: *MP2CY*
Source: Performance data set
Values: Integers ranging from 23 to 3,384
Coding: How many minutes of playing time a player logged two years prior to contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mp2cy	1441.141	862.7512	717	1388.5	2167	23	3384

Histogram:

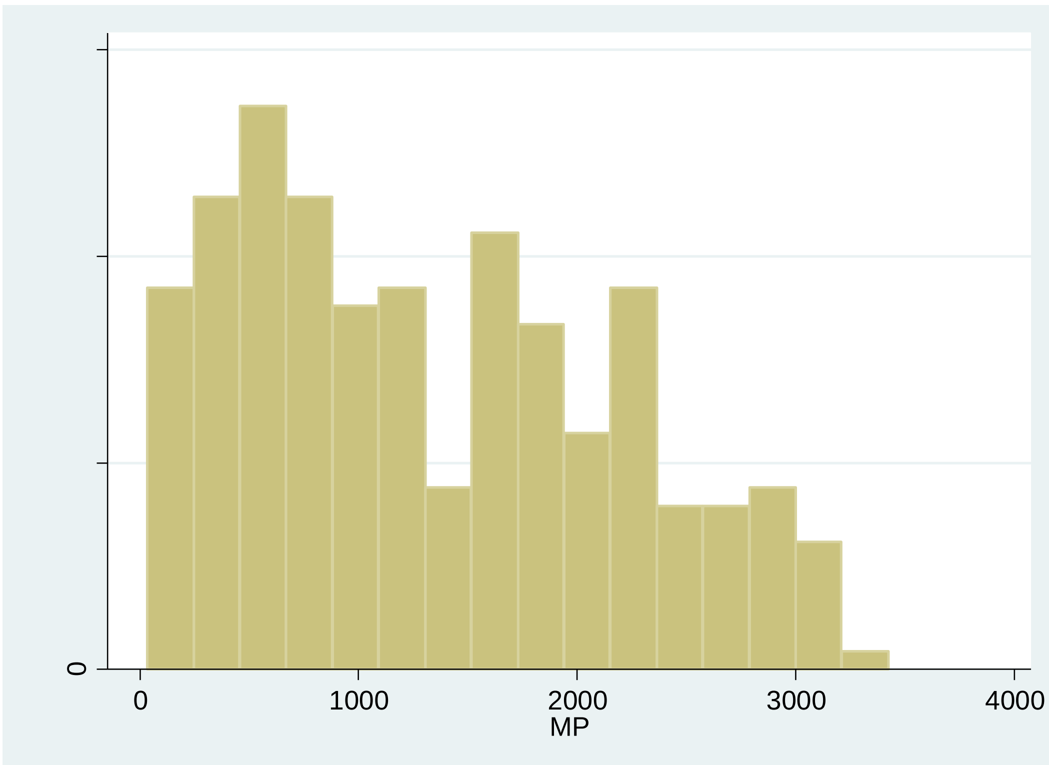


Variable name: *MP1CY*
Source: Performance data set
Values: Integers ranging from 34 to 3,424
Coding: How many minutes of playing time a player logged one year prior to contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mp1cy	1330.28	856.056	590.5	1175.5	1977.5	34	3424

Histogram:

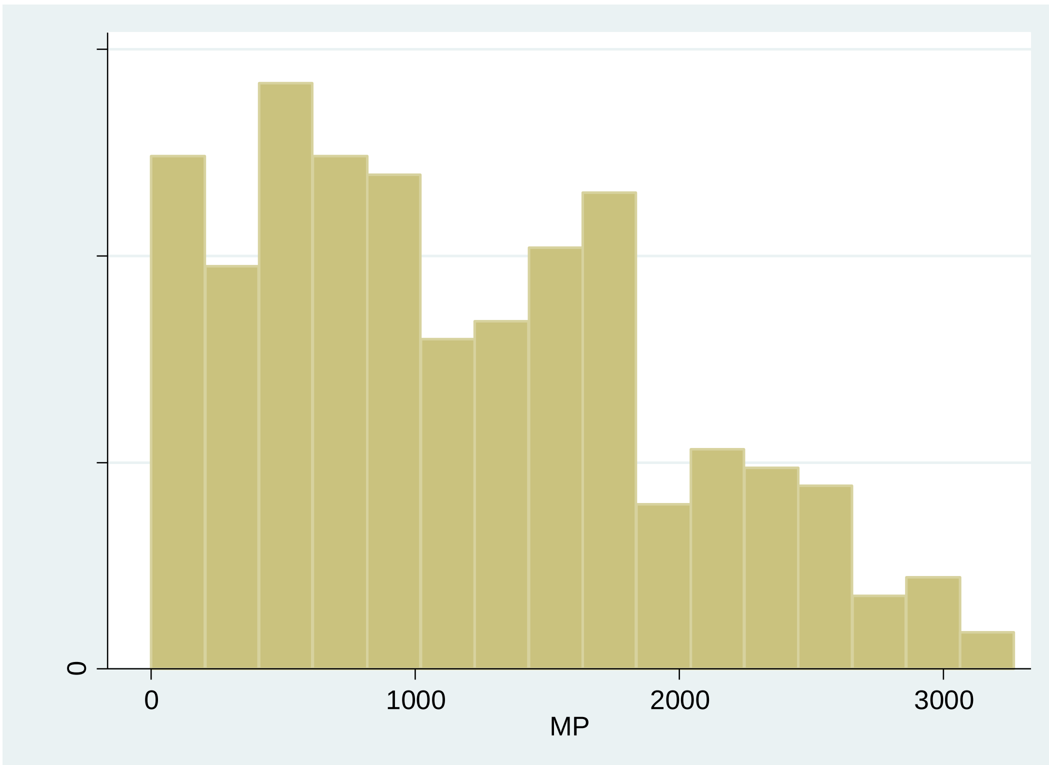


Variable name: *MPCY1*
Source: Performance data set
Values: Integers ranging from 0 to 3,269
Coding: How many minutes of playing time a player logged one year after contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mpcy1	1167.978	780.2741	513.5	1032	1687.5	0	3269

Histogram:

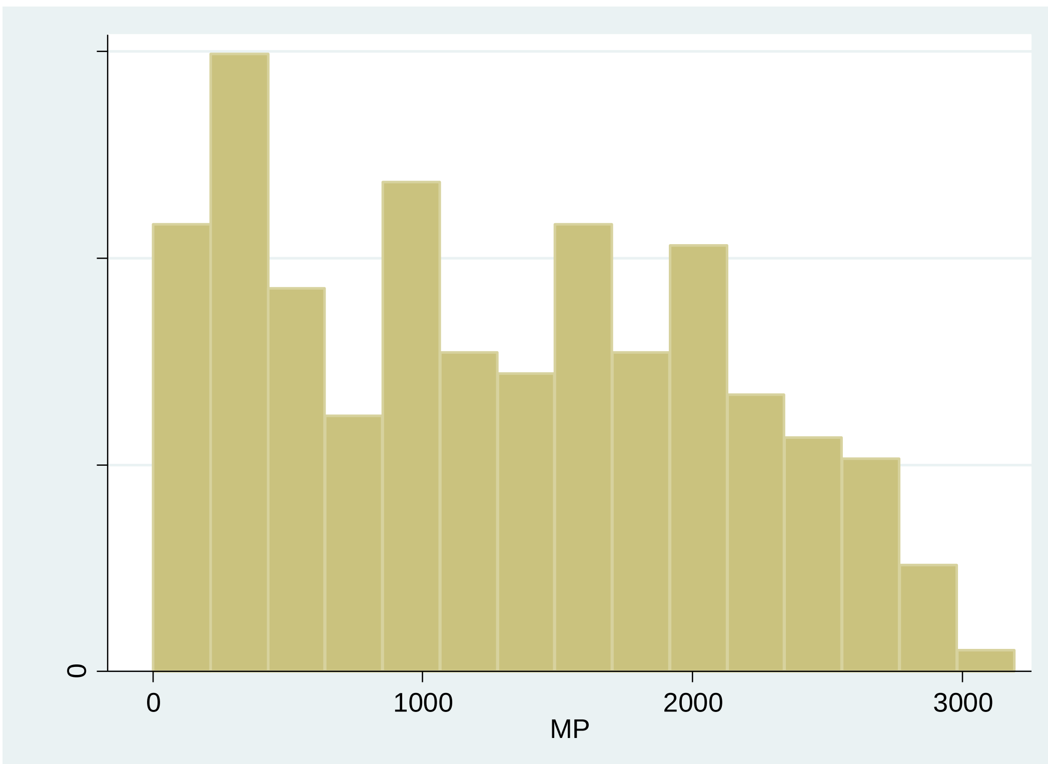


Variable name: *MPCY2*
Source: Performance data set
Values: Integers ranging from 1 to 3,193
Coding: How many minutes of playing time a player logged two years after contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mpcy2	1268.167	819.7114	497	1192.5	1941	1	3193

Histogram:

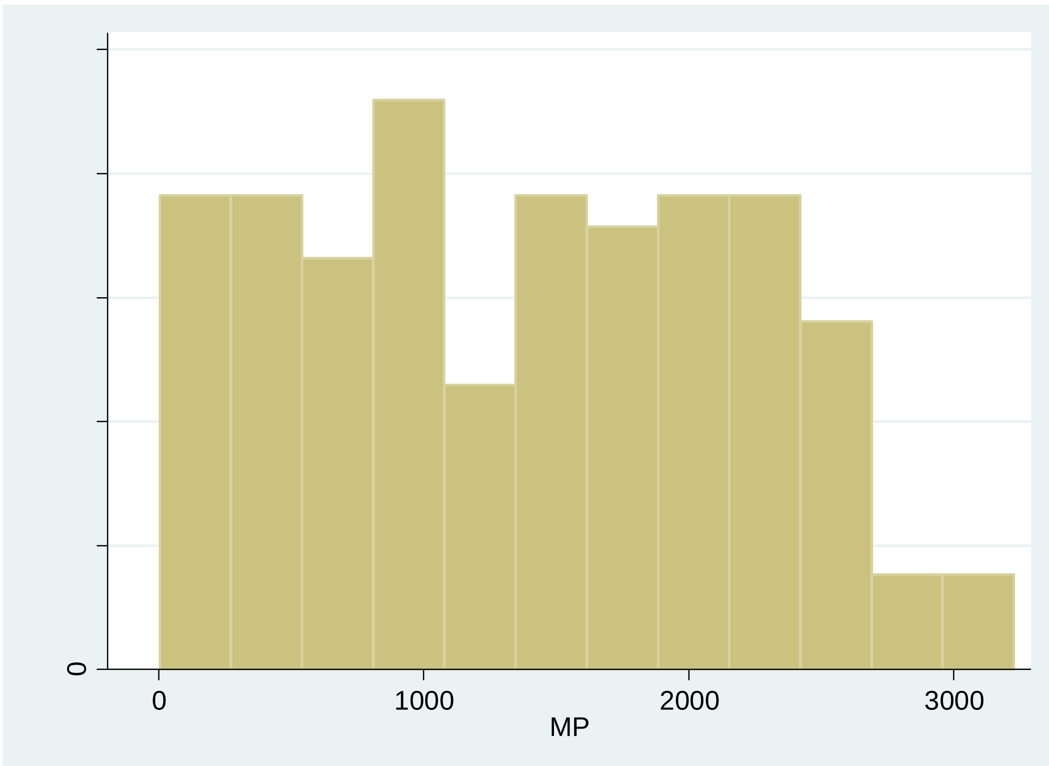


Variable name: *MPCY3*
Source: Performance data set
Values: Integers ranging from 5 to 3,227
Coding: How many minutes of playing time a player logged three years after contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mpcy3	1382.904	818.5867	694	1388	2100	5	3227

Histogram:

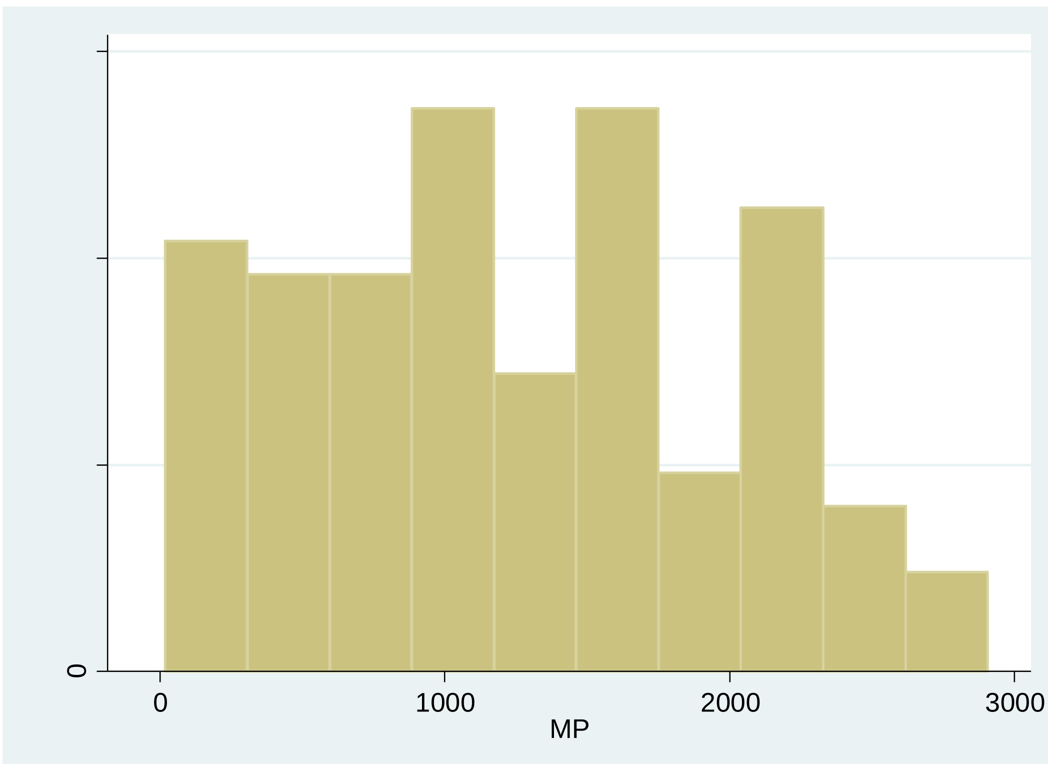


Variable name: *MPCY4*
Source: Performance data set
Values: Integers ranging from 18 to 2,907
Coding: How many minutes of playing time a player logged four years after contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mpcy4	1263.685	745.7467	676.5	1192	1793.5	18	2907

Histogram:

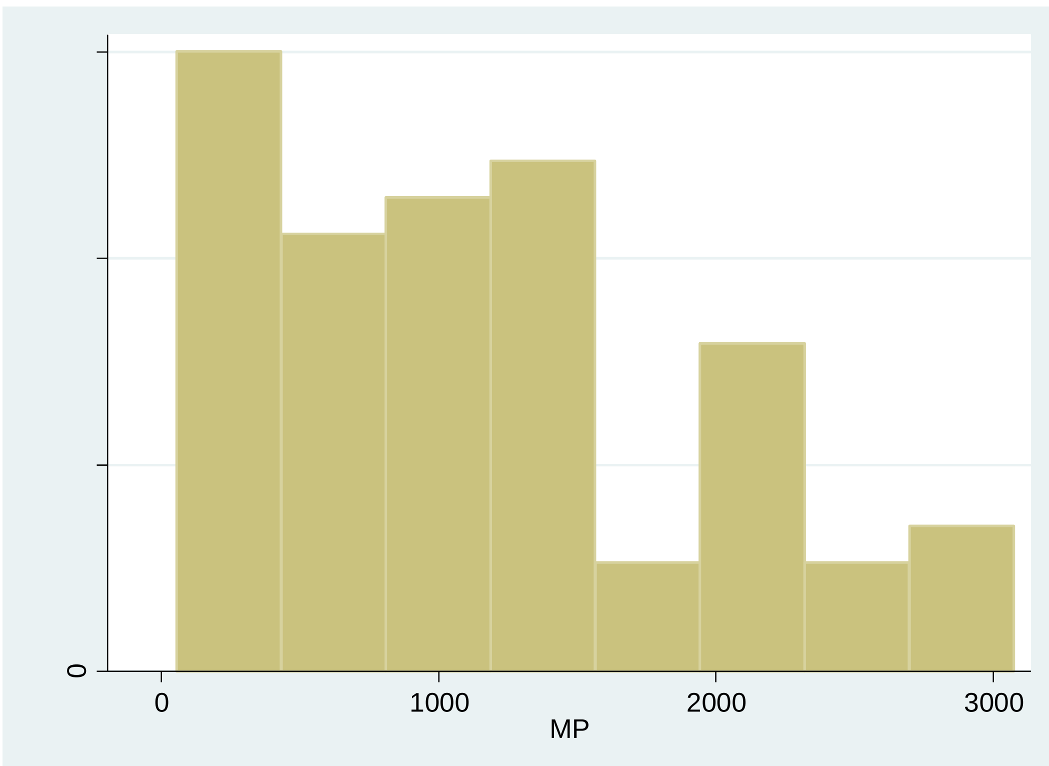


Variable name: *MPCY5*
Source: Performance data set
Values: Integers ranging from 55 to 3,076
Coding: How many minutes of playing time a player logged five years after contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
mpcy5	1176.827	789.2845	482	1039	1674	55	3076

Histogram:

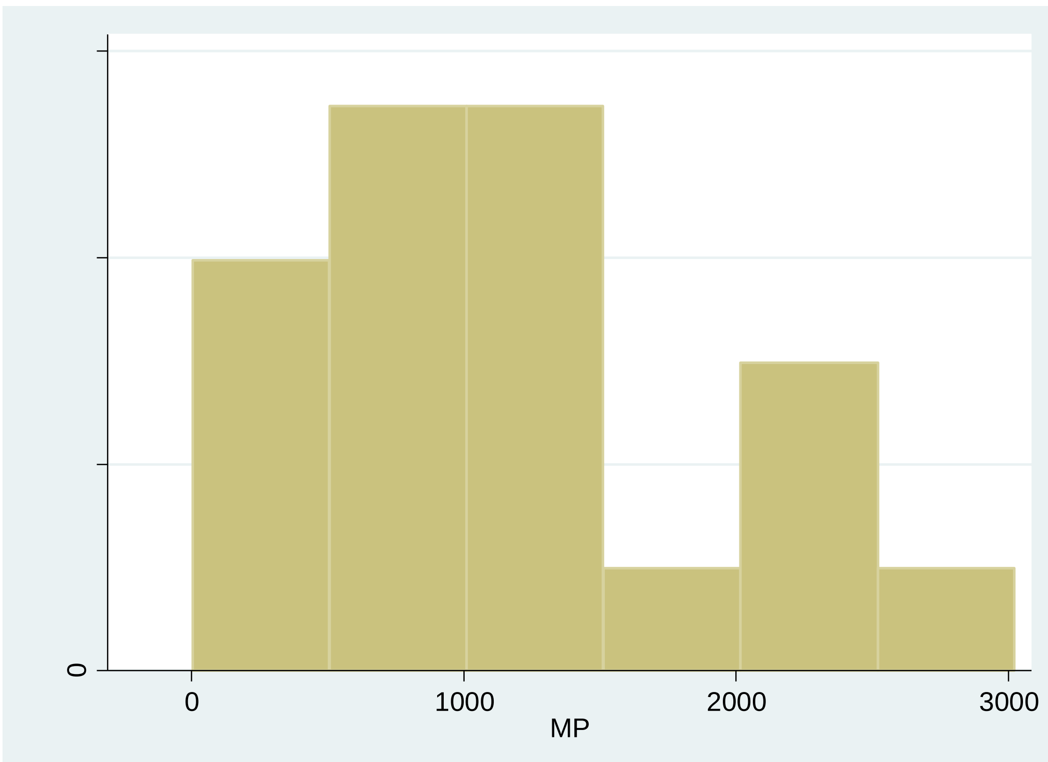


Variable name: *MPCY6*
Source: Performance data set
Values: Integers ranging from 5 to 3,023
Coding: How many minutes of playing time a player logged six years after contract year

Summary Statistics:

```
    variable |       mean         sd        p25        p50        p75        min        max
-------------+-----------------------------------------------------------------------------
       mpcy6 |    1174.75    766.081        536     1166.5       1513          5       3023
-------------------------------------------------------------------------------------------
```

Histogram:

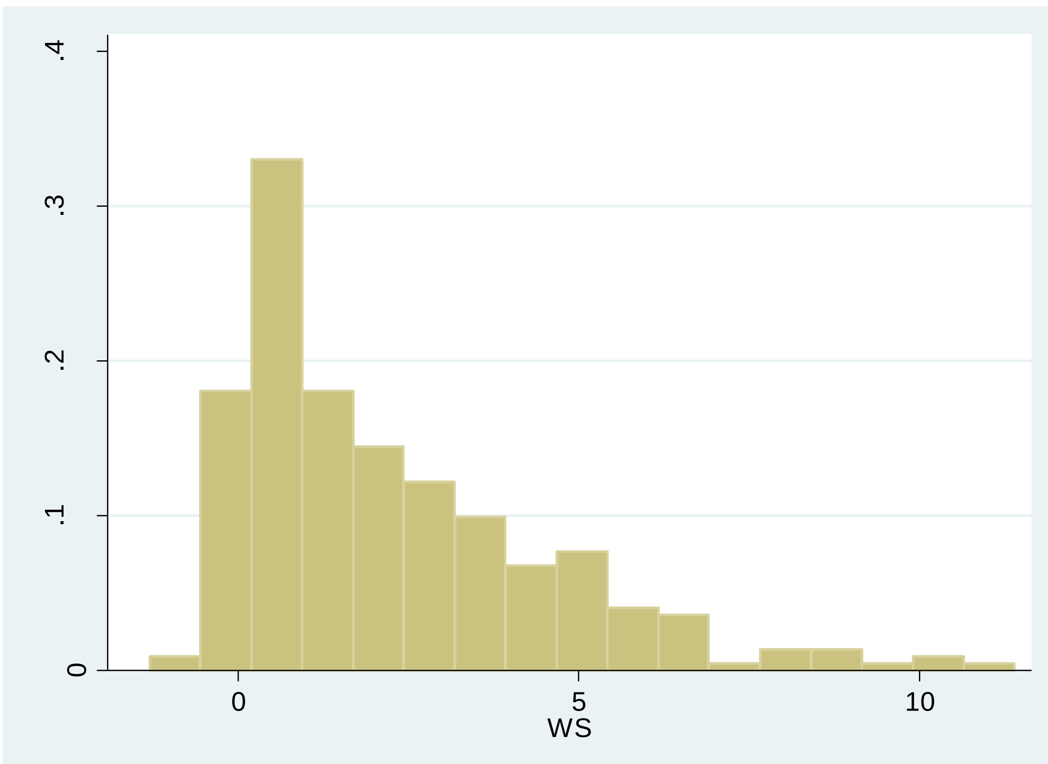


Variable name: *WSCY*
Source: Performance data set
Values: Numbers ranging from -1.3 to 11.4
Coding: The win share value calculated for each player during their contract year season

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy	2.226351	2.285914	.5	1.5	3.45	-1.3	11.4

Histogram:

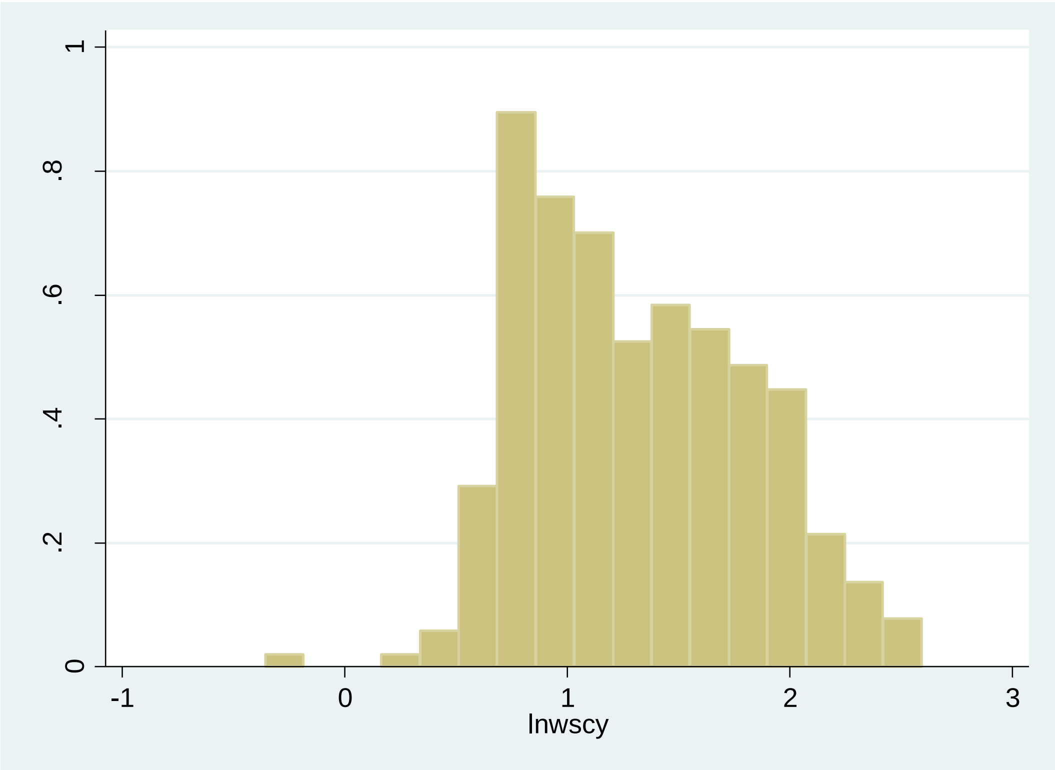


Variable name: *LNWSCY*
Source: Performance data set
Values: Numbers ranging from -0.35 to 2.60
Coding: The natural log of the win share value calculated for each player during their contract year season plus 2 (to keep all numbers feasible for natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy	1.312029	.5028744	.9162908	1.252763	1.695574	-.3566749	2.595255

Histogram:

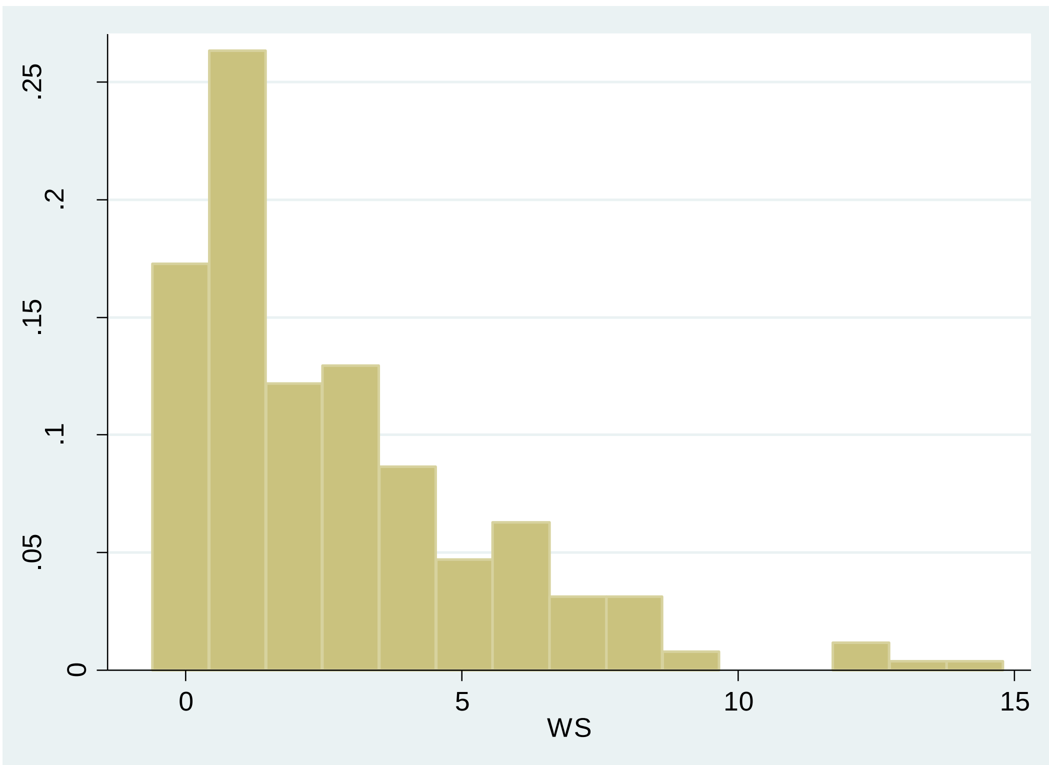


Variable name: *WS2CY*
Source: Performance data set
Values: Numbers ranging from -0.6 to 14.8
Coding: The win share value calculated for each player during the season two years prior to their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
ws2cy	2.725	2.734208	.7	1.9	4.15	-.6	14.8

Histogram:

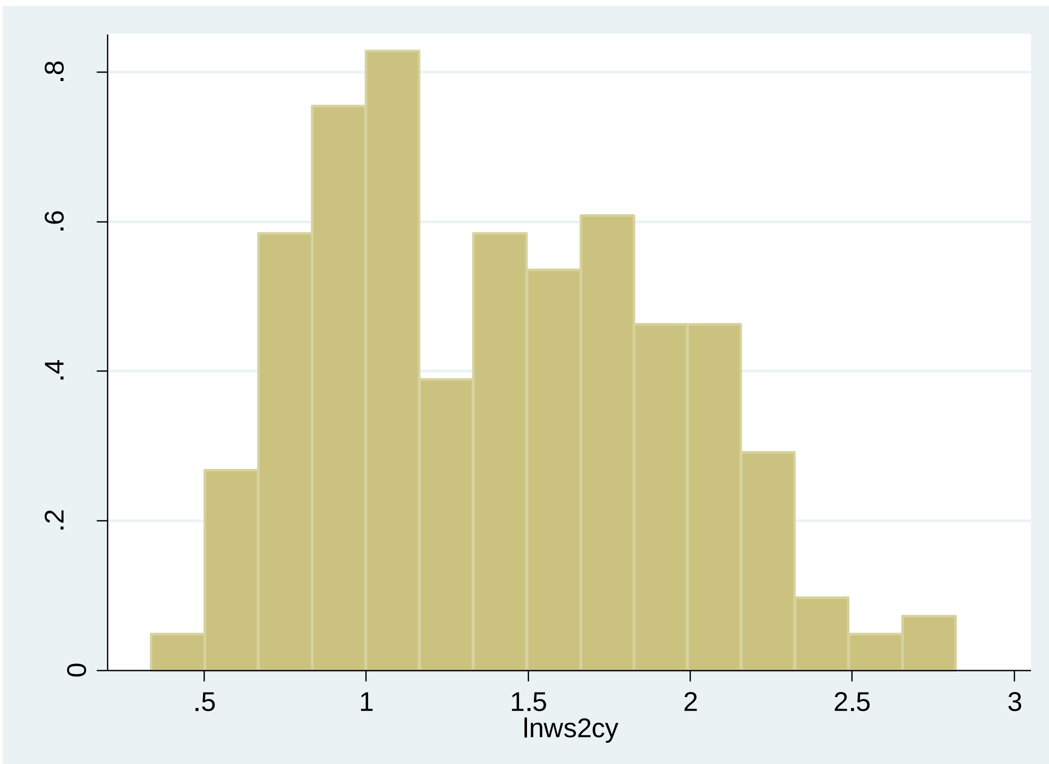


Variable name: *LNWS2CY*
Source: Performance data set
Values: Numbers ranging from .33 to 2.83
Coding: the natural log of the win share value calculated for each player during the season two years prior to their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnws2cy	1.409378	.5263981	.9932518	1.360977	1.816419	.3364722	2.821379

Histogram:

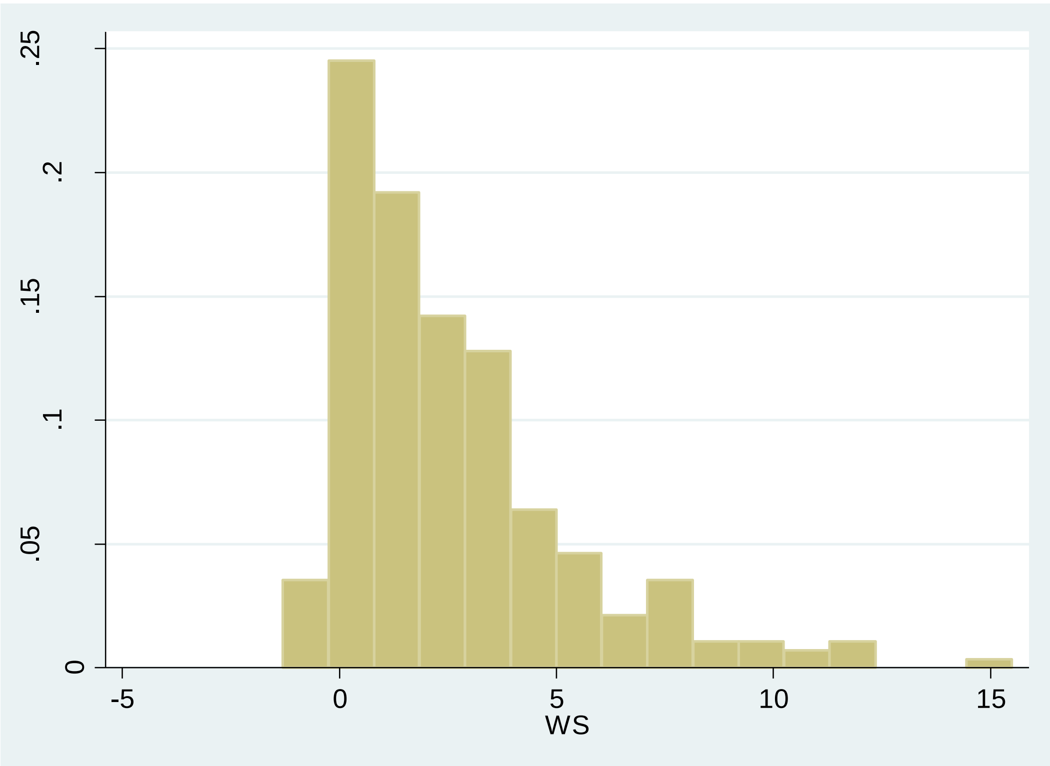


Variable name: *WS1CY*
Source: Performance data set
Values: Numbers ranging from -1.3 to 15.5
Coding: The win share value calculated for each player during the season one year prior to their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
ws1cy	2.561194	2.667249	.6	1.9	3.6	-1.3	15.5

Histogram:

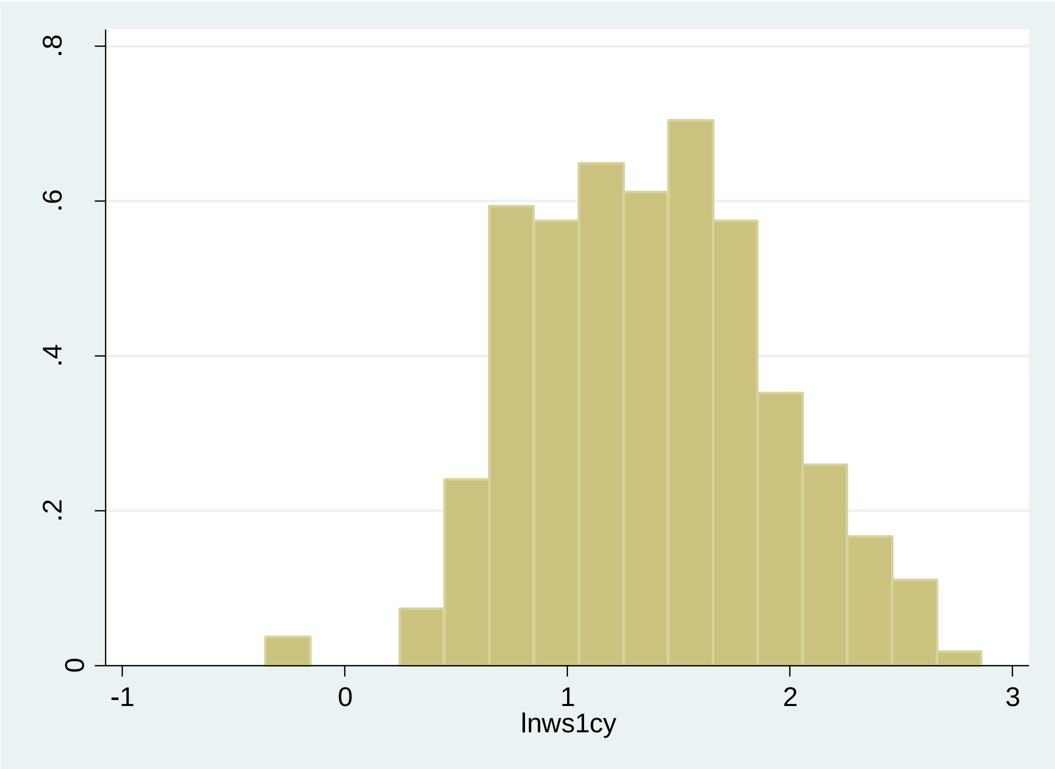


Variable name: *LNWS1CY*
Source: Performance data set
Values: Numbers ranging from -0.36 to 2.87
Coding: the natural log of the win share value calculated for each player during the season one year prior to their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnws1cy	1.369482	.540881	.9555115	1.360977	1.722767	-.3566749	2.862201

Histogram:



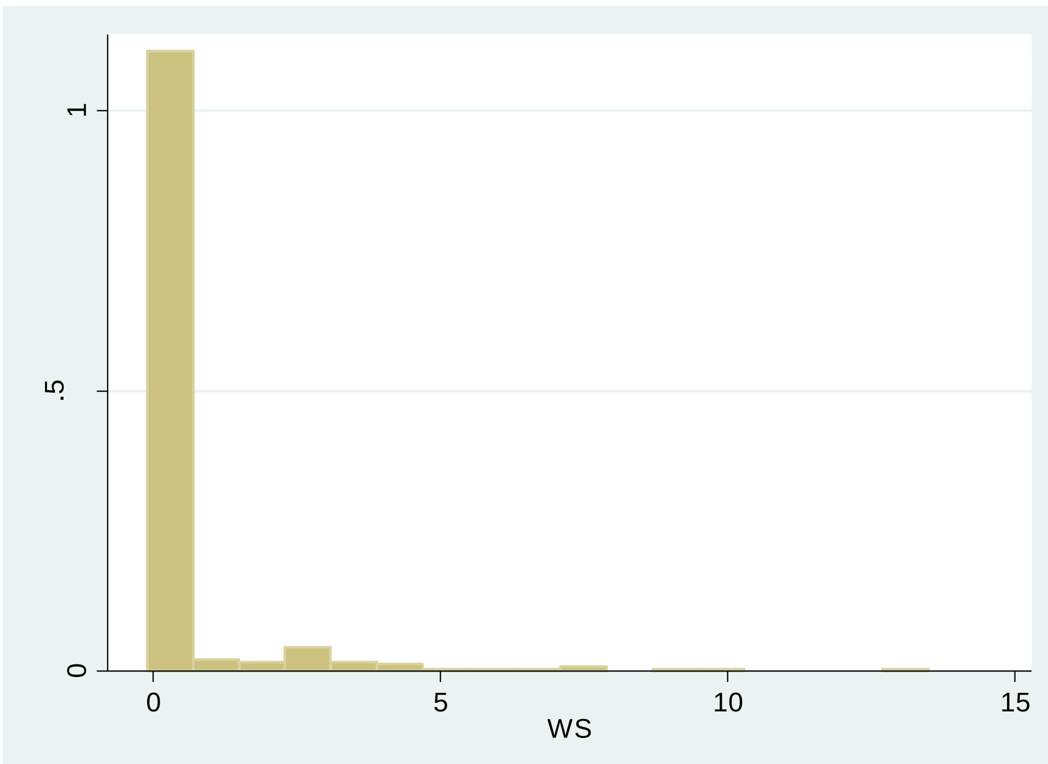
Variable name: *WSCY1*
Source: Performance data set
Values: Numbers ranging from -0.1 to 13.5
Coding: the win share value calculated for each player during the season one year after their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy1	.4486486	1.566643	0	0	0	-.1	13.5

H

Histogram:

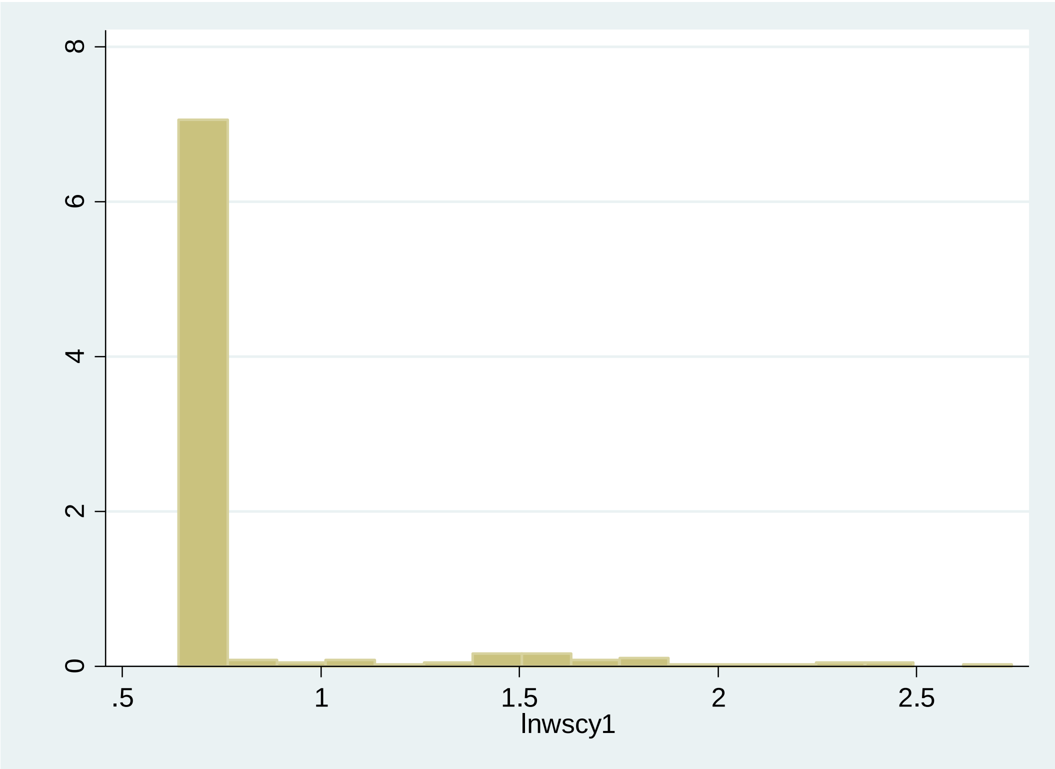


Variable name: *LNWSCY1*
Source: Performance data set
Values: Numbers ranging from 0.69 to 2.75
Coding: the natural log of the win share value calculated for each player during the season one year after their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy1	.8076598	.344595	.6931472	.6931472	.6931472	.6418539	2.74084

Histogram:

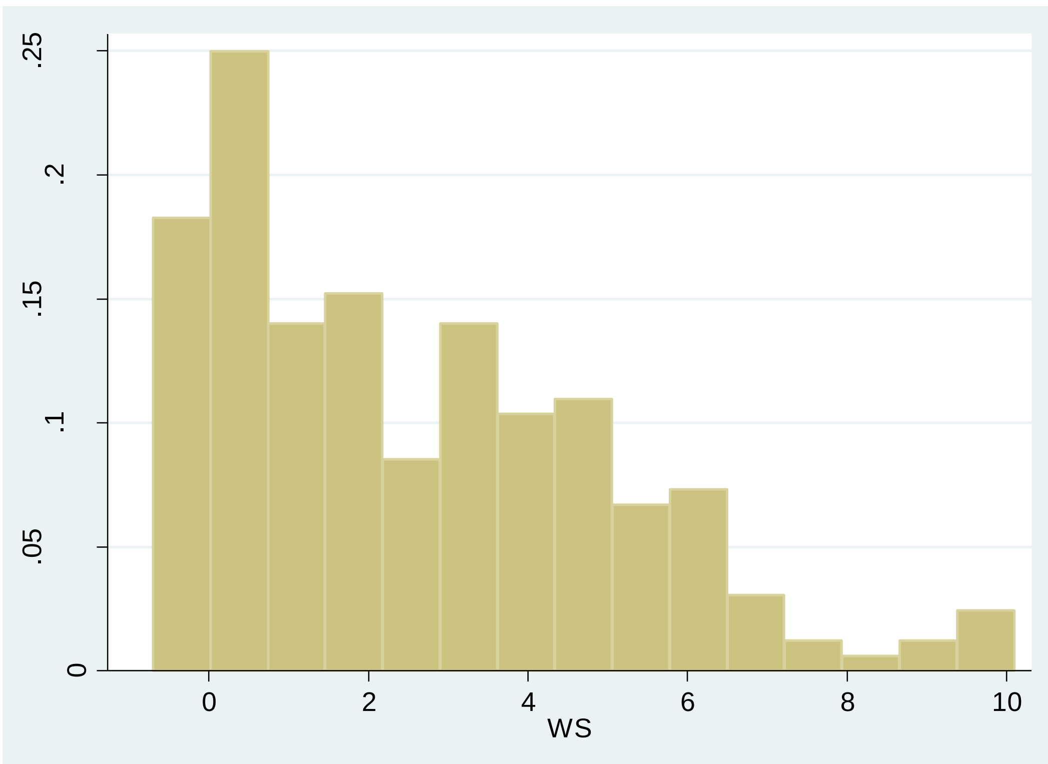


Variable name: *WSCY2*
Source: Performance data set
Values: Numbers ranging from -.7 to 10.1
Coding: the win share value calculated for each player during the season two years after their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy2	2.577632	2.414052	.4	2	4.3	-.7	10.1

Histogram:

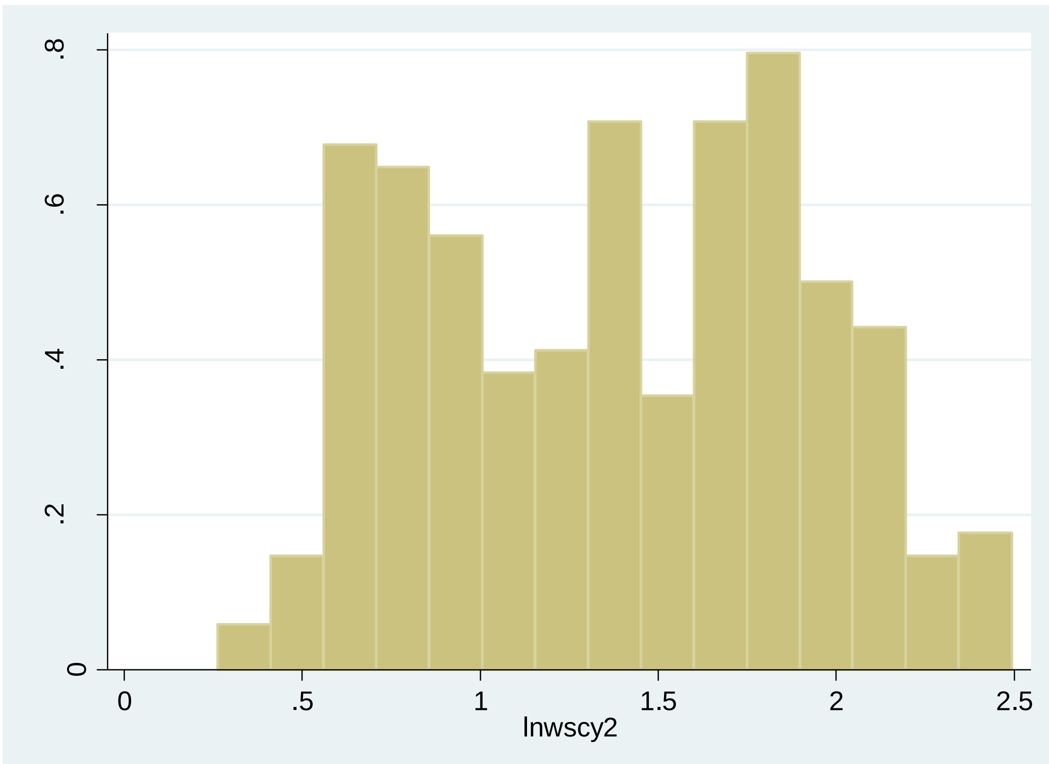


Variable name: *LNWSCY2*
Source: Performance data set
Values: Numbers ranging from .26 to 2.50
Coding: the natural log of the win share value calculated for each player during the season two years after their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy2	1.383506	.5310861	.8754687	1.386294	1.84055	.2623643	2.493206

Histogram:

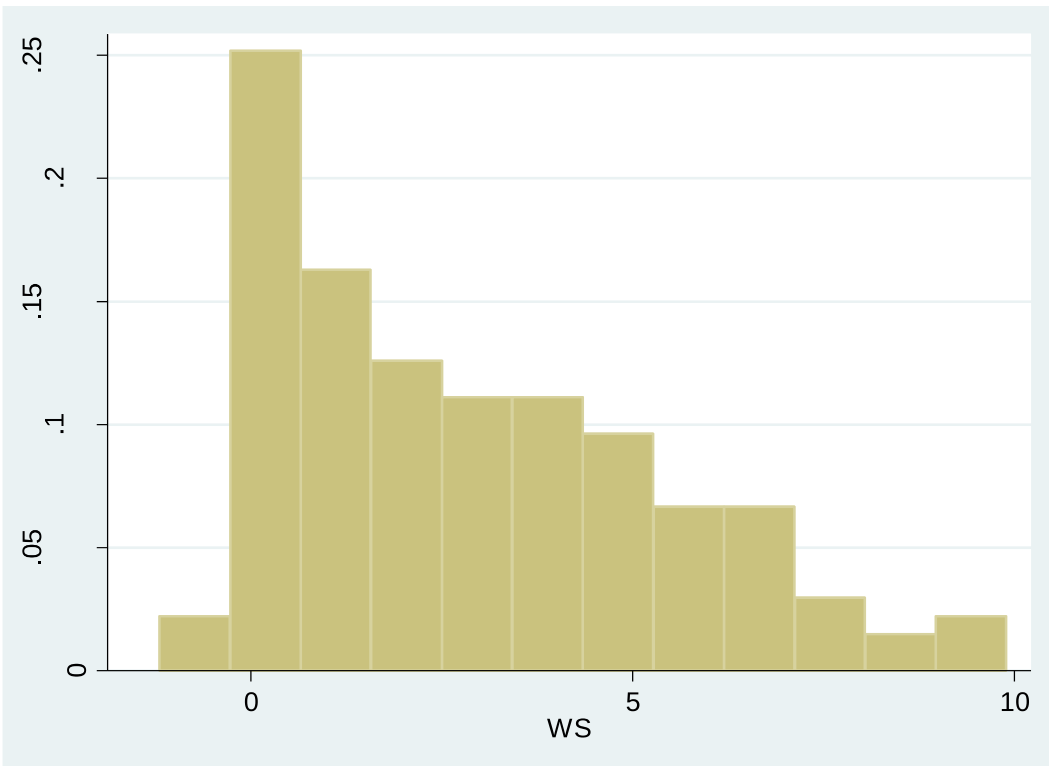


Variable name: *WSCY3*
Source: Performance data set
Values: Numbers ranging from -1.2 to 9.9
Coding: the win share value calculated for each player during the season three years after their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy3	2.828082	2.4918	.6	2.35	4.7	-1.2	9.9

Histogram:

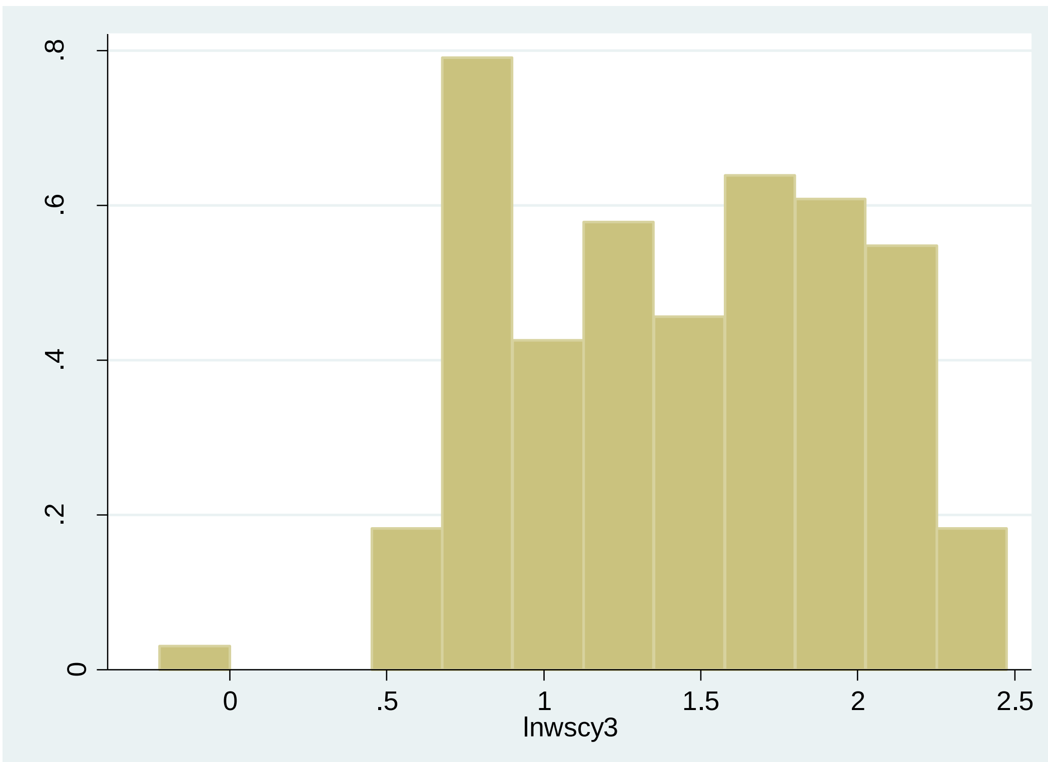


Variable name: *LNWSCY3*
Source: Performance data set
Values: Numbers ranging from -0.22 to 2.48
Coding: the natural log of the win share value calculated for each player during the season three years after their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy3	1.43734	.5370463	.9555115	1.47011	1.902107	-.2231435	2.476538

Histogram:

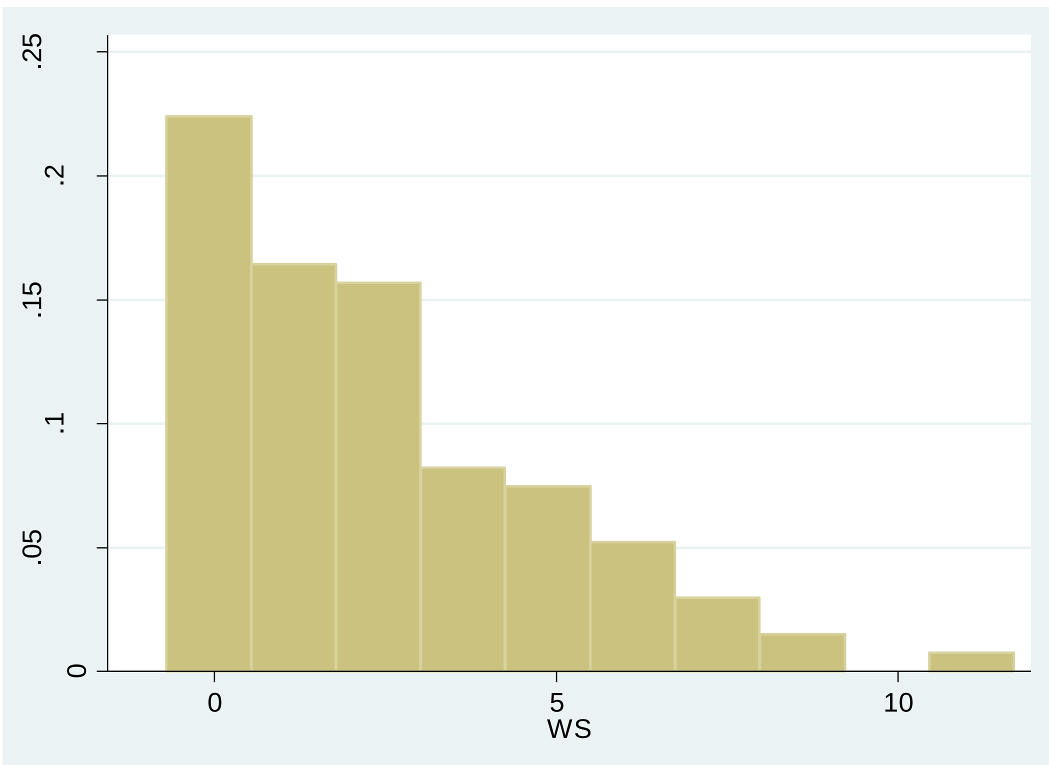


Variable name: *WSCY4*
Source: Performance data set
Values: Numbers ranging from -0.7 to 11.7
Coding: the win share value calculated for each player during the season four years after their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy4	2.486111	2.411439	.45	1.8	4.1	-.7	11.7

Histogram:

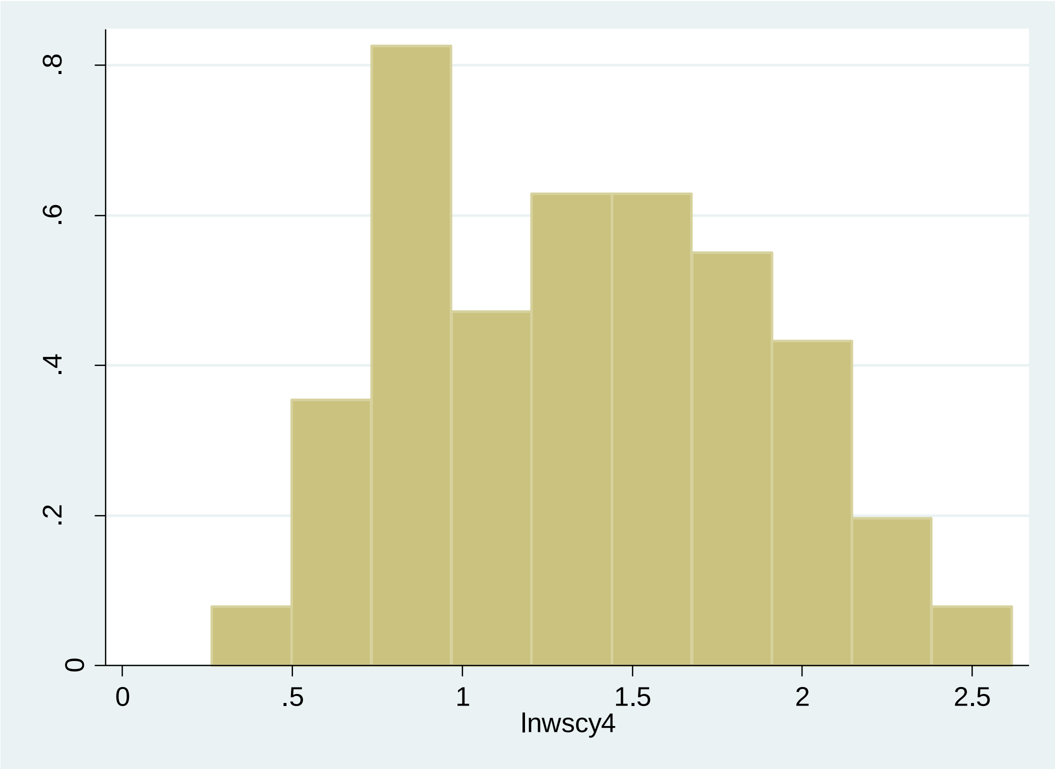


Variable name: *LNWSCY4*
Source: Performance data set
Values: Numbers ranging from 0.26 to 2.62
Coding: the natural log of the win share value calculated for each player during the season four years after their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy4	1.366262	.5213914	.8958797	1.335001	1.808154	.2623643	2.617396

Histogram:

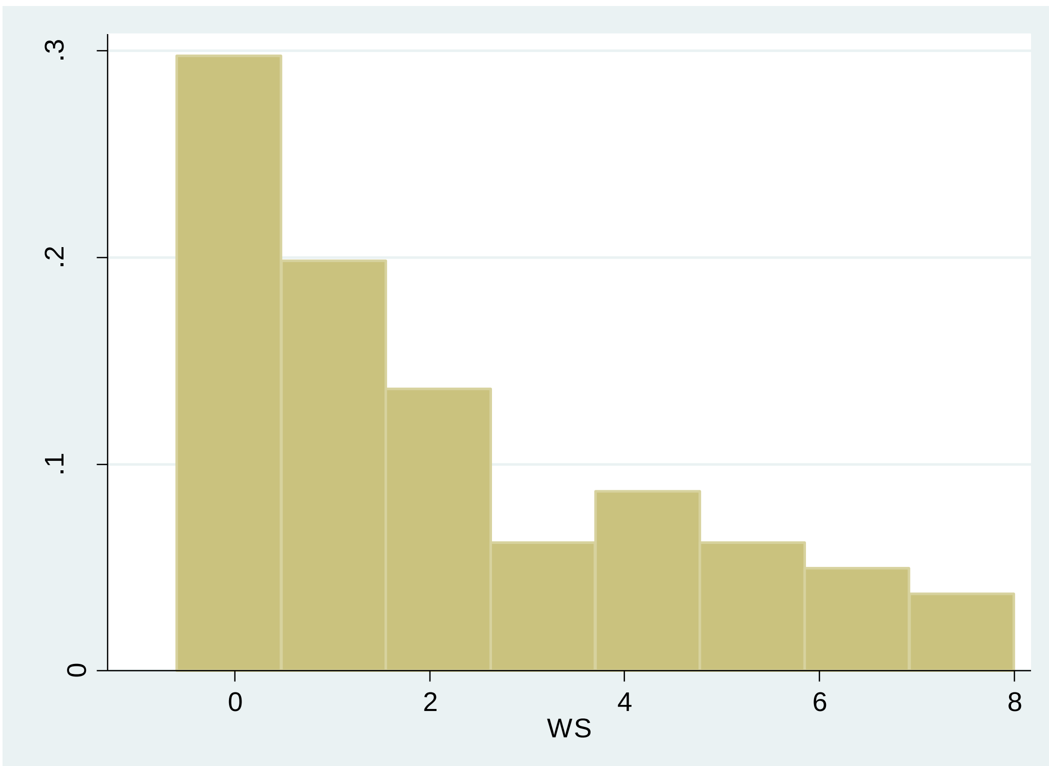


Variable name: *WSCY5*
Source: Performance data set
Values: Numbers ranging from -0.6 to 8
Coding: the win share value calculated for each player during the season five years after their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy5	2.113333	2.271643	.2	1.4	4	-.6	8

Histogram:

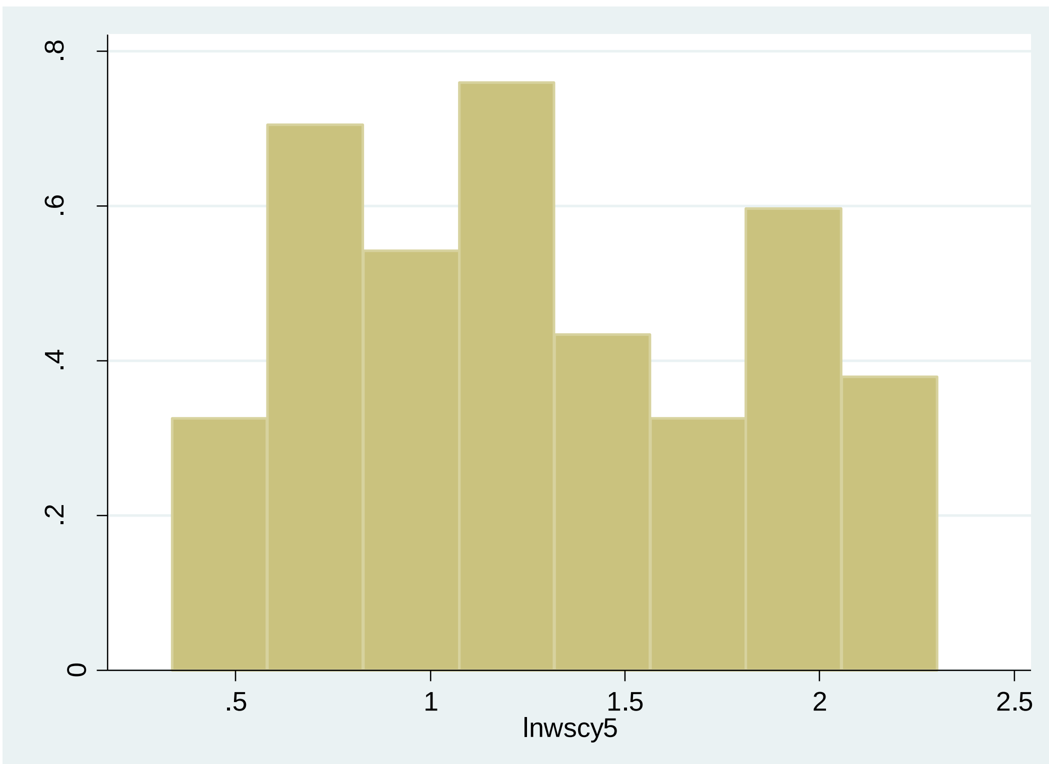


Variable name: *LNWSCY5*
Source: Performance data set
Values: Numbers ranging from 0.33 to 2.31
Coding: the natural log of the win share value calculated for each player during the season five years after their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy5	1.271815	.5342271	.7884573	1.223775	1.791759	.3364722	2.302585

Histogram:

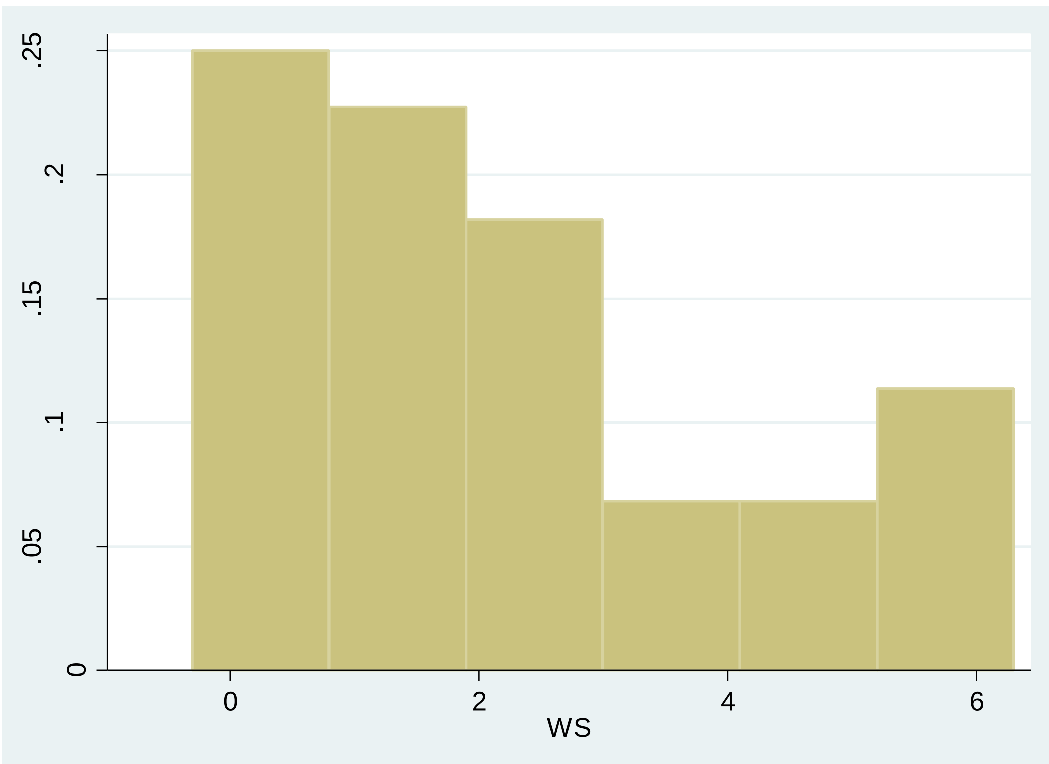


Variable name: *WSCY6*
Source: Performance data set
Values: Numbers ranging from -0.3 to 6.3
Coding: the win share value calculated for each player during the season six years after their contract year

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
wscy6	2.18	2.058404	.55	1.65	3.4	-.3	6.3

Histogram:

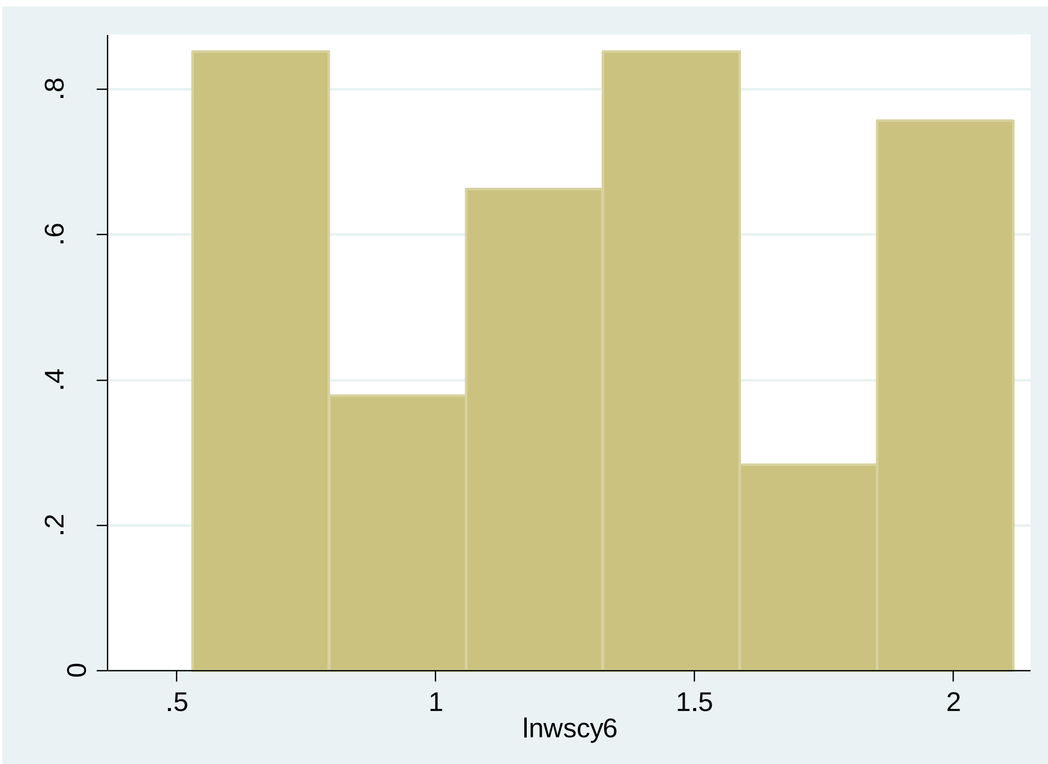


Variable name: *LNWSCY6*
Source: Performance data set
Values: Numbers ranging from 0.53 to 2.12
Coding: the natural log of the win share value calculated for each player during the season six years after their contract year plus two (to keep all numbers feasible for taking the natural log)

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lnwscy6	1.312675	.4945382	.9343603	1.293882	1.686227	.5306283	2.116256

Histogram:



Variable name: *LNTOTFUTWS*
Source: Performance data set
Values: Numbers ranging from 0.69 to 13.06
Coding: The sum of the natural log of the win share values calculated for each player during each of the seasons during their free agent contract

Summary Statistics:

variable	mean	sd	p25	p50	p75	min	max
lntotfutws	2.220025	2.440381	.6931472	.6931472	2.555656	.6931472	13.05485

Histogram: