Simplicity Versus WAR:

Examining Salary Determinations in Major League Baseball's Arbitration and Free Agent Markets

Joshua Studnitzer

Richard Ball, Ph.D., Advisor

Senior Thesis in Economics

Haverford College

May 1, 2014

Abstract

This paper examines salaries given to arbitration eligible players in Major League Baseball from 2008-2013 and compares them to free agent contracts from the same period. Anecdotal evidence suggests that simpler statistics are more successful in Major League Baseball's final offer arbitration setting as legal experts tasked with handling the league's cases may not have a deep knowledge of player valuation. I examine the effects of wins above replacement, a complex but comprehensive metric, and traditional statistics, such as runs batted in, on salaries decided in both settings. Wins above replacement is significant in each case, but with a much higher coefficient in free agency suggesting a greater impact. There is no evidence of individual traditional statistics being especially significant in arbitration; I attribute this to parties framing their offers with whichever statistics portray them in the most favorable light. Finally, I look to statistics in the season following contracts to determine if either market is more effective in getting value at a low cost, but results are similar in each case and limitations with the data restrict the efficacy of conclusions in the section.

Table of Contents

1. Introduction and Background

Over the past decades, Major League Baseball has vaulted into its information age. Teams utilize cutting-edge data gathering systems, hire employees with advanced degrees from America's finest institutions, and display a constant thirst for innovation. Among the many developments of this transition has been the modernization of player valuation as teams scrutinize a player's talents on the most granular of levels. Using this information, they can come to a comprehensive assessment of the player's previous production, which, combined with age, provides a good sense of expected future production. These estimations are commonplace in decisions regarding free agent players, whom each team is able to carefully examine and offer a salary (or not) based on their potential contributions over the contract length.

As these metrics were beginning to gain favor, near-systemic misevaluation was common and exploitable. Fast-adapting teams encountered vast advantage as the correlation between player production, now more accurately calculated, and salary on the open market were misaligned. Moneyball, a bestseller by Michael Lewis, discusses on base percentage, situating it as the most familiar source of misalignment. However, others factors like defense and contextual information, such as a player's home ballpark, leading to inflated or diminished statistics also were misjudged and led to undervalued players. These misalignments were fleeting, however, as most teams caught on, quickly removing the widespread potential for surplus value in the free agent market. My goal is to examine another market, arbitration, and determine whether proper valuation is present or if the opportunity to take advantage persists.

Wins above replacement, referred to from here on as WAR, is the metric I choose to quantify a player's absolute production in a way that mimics the process of evaluation occurring in front offices. The version I use is taken from Baseball-Reference.com and I employ it as a proxy for these assessments within my regressions. Assuming it is successful in capturing a front office's outlook, WAR would have a large and highly significant coefficient on contract size in free agency.

It is important to discuss the process of arbitration, which may structurally encourage an emphasis away from these more accurate, yet complicated, metrics. A team maintains exclusive rights to a player's talents for his first six years of service time in the Major Leagues. After three years of monopsonostic team control, a player has three years in which he can choose to file for arbitration to determine his salary. Major League Baseball utilizes a final offer arbitration system to determine these contracts. Essentially, this means that an arbitrator must choose between salaries submitted by the player's camp and the team's baseball operations department without the opportunity to find a middle ground. The panels are composed of esteemed arbitrators with extensive legal experience; however, they do not necessarily have equal depth of knowledge with respect to baseball, especially in terms of groundbreaking player valuation. Hearings are structured so that each side receives ninety minutes to present its case (an hour to present and thirty minutes to rebut and for summation), and arbitrators are not informed of their cases beforehand so they are unable to prepare. The confluence of these factors promotes simplicity. Both players and teams are incentivized to promote their case using the simplest, most accessible argument, to convince the arbitrator with limited exposure of the validity of their salary figure.

Anecdotal evidence suggests that involved parties are more successful in arbitration when framing their arguments with basic traditional statistics. John Copolella, an assistant general manager for the Atlanta Braves who is heavily involved in the team's arbitration dealings, says, "Many of the arguments put forth by both sides are predicated on counting statistics like home runs, runs batted in, and runs scored. Ratio statistics like OPS {on base plus slugging] or wOBA {weighted on base average} are used, but tend to carry less weight"(Coppolella 1). Gregg Clifton, a prominent player agent similarly states, "You have to figure out what is going to be the best information to win your case in a one-hour presentation. … We can't assume … that all the arbitrators are going to understand all these advanced analytics. We also don't have the 10 minutes to explain the analytic and explain why it's important and why it's relevant to our case, because now we've just lost 10 of our 60 minutes" (Clifton 1). Roger Abrams provides evidence from the perspective of an arbitrator with experience in Major League Baseball cases in his book The Money Pitch. "An agent who does well in salary arbitration focuses on the core characteristics of his player. A team wins by scoring runs. Run production – runs scored and RBIs – is the key offensive statistic"(Abrams 158). It is clear that the perception exists within arbitration that simplicity is rewarded; in the pages that follow, I will empirically examine this view.

I expect to find that WAR has a larger impact on salary in free agency than in arbitration. However, I am also interested in examining whether simple traditional statistics individually are significantly more influential in arbitration as Copolella argues. For example, runs batted in offers a measure that has a lofty place in baseball's tradition, but has been downplayed as analytics have proliferated (since it is determined more by a player's context in the lineup than a player's skill). It is possible that arbitrators could resort to numbers such as these as a crutch, but I do not expect this to be the case. Rather, my hypothesis would be that participants in arbitration would construct a cogent but accessible argument framed with whatever would advance their side's case most effectively. The goal is to convince the arbitrator of a player's inflated or deflated worth so discussing the greatest flaws or attributes in a player's skillset would be the most useful strategy. This could be any number of factors, and thus would not show up in the effect of any specific statistic, but it would mean that WAR would be deemphasized.

To reiterate my two main hypotheses, I expect prior production as determined by WAR to be more significant in determining salaries in free agency than in arbitration. Secondly, I do not expect there to be other predictors that are largely significant in the case of arbitration since involved parties will frame their cases using anything that provides the most positive basic argument of their case.

Ultimately, I find that WAR had a significant positive coefficient in each market, but the coefficient was much larger in free agency, meaning that it plays a greater role in those negotiations. Meanwhile, there was no significant difference between the effect of traditional statistics in free agency and arbitration, suggesting that there is no statistic that consistently drives the decision of arbitrators.

2. Literature Review

Fleshing out the motivations and reasoning of each party involved in an arbitration hearing is crucial in understanding the resulting salaries. Marburger (2004) examines the chilling effect within final offer arbitration. The chilling effect argues that an arbitrator will find a happy medium between two offers in a conventional arbitration system and is ostensibly removed by the requirement that the arbitrator chooses only between the two submitted bids. However, Marburger argues that baseball's final offer system still generates compromise. as each party tends to moderate its offer to increase its chances of winning. Thus, the offers exchanged to the arbitrators are likely to slightly understate the perceived value in the player's case and overstate it from the team's perspective. In terms of a salary determined, this should produce a result similar to free agency in that the ultimate salary falls, on average, more or less at the midpoint of the two estimates. This number should be approximately equal to free agency on average, except for the role of the arbitrator. Therefore, it seems to be a safe conclusion that distinctions in terms of drivers seen in the two subsets in my regressions can in fact be attributed to the arbitrator.

In modeling the arbitration process, Miller (2000) allows for a two-step approach that incorporates the potential for settlement in the days between the sharing of offers and the arbitration hearing. The degree to which each party disagrees with the offer of the other party becomes relevant in this model as they become less likely to come to terms on an agreement with large disagreement outcomes or heavy negotiating costs. As he puts it "in the face of high negotiating costs, the negotiators may find it optimal to forgo bargaining and instead opt for an arbitrated solution" (Miller 92). Negotiating costs will be heightened if a party has an improved chance of winning a trial, and if there is a player whose value would be predictably miscalculated by an arbitrator by virtue of a stat like his number of runs batted in, the costs jump for the side that is favored. Therefore, even if cases were settled before the trial begins, as long as all parties are aware of arbitrator bias, any inefficiency would be reflected in salaries. This is justification of my decision to include salaries for all arbitration-eligible players in my dataset, which I will cover in the "Data" section.

Fizel (2002) examines equity as a potential important aspect of the arbitration process. If equity did influence a player's willingness to settle, it would be deleterious to my conclusions as rationality could be succumbing to stubbornness in the player's refusal to accept an offer as a result of perceived past unjust compensation. However, he finds that while equity is important in the player's initial decision to file, it is not a factor in the chance that the player goes to a hearing. This is logical since the team often waits to conduct negotiations until the player has filed.

Burger (2005) finds that arbitration earnings as a percentage of marginal revenue product increase as a player gains years of experience. This indicates that the market in arbitration acts as a bridge between pre-arbitration and free agency. One explanation is that a player must make at least 80 percent of their previous salary in an arbitration hearing (and players at most levels of production will see an increase). Therefore, the process engenders a steady escalation of salary throughout the three years. Since the arbitrators are aware that the first year arbitration-eligible players will likely be seeing raises in the two years that follow, perhaps they are compensating for this by under assessing the first years. The does suggest that special attention should be given to a player's class of service time in each of my regressions, which I include as an independent variable, along with service time squared, which allows for the parabolic effect that age-related decline likely creates.

Miller (2000) offers a similar study to mine, conducting an empirical comparison between free agent and arbitration-determined salaries. He runs two separate regressions and then conducts a Chow test for equality of salary structures. In both regressions, previous salary, runs created and inverse of games played are highly significant with runs created and previous salary being positive in both. Games played are positive in arbitration and negative in free agency, a likely effect of wear and tear of older free agents. Middle of the diamond players (catchers, shortstop, second basemen, and center fielders) make significantly more only in arbitration. National League/American League, defensive metrics, and winning percentage are all insignificant. This is a empirical indication that distinctions occur, and I hope to update his work, albeit looking at different factors, as the composition of front offices involved in negotiations have been revolutionized and more fine-tuned evaluative measures are easily accessible to investigate outcomes.

3. Data

For arbitration eligible players, I was able to compile my own dataset including every player who has filed for arbitration in the off-seasons preceding 2008 to 2013. This list of players was available on www.bizofbaseball.net and included player, team, position, service time, figures submitted by the player and club, ultimate salary, and winning party if applicable. To this information, I merged hitting statistics queried from the Play Index on Baseball-Reference.com. These include WAR and a number of traditional stats, such as runs batted in and home runs. Free agency contract terms are available dating back to 2008 on ESPN.com and I appended them to my arbitration dataset, marking them with a dummy variable to separate the observations.

It is important to establish that I am looking beyond just the players who reach an arbitration hearing in conclusions I hope to make. Disincentives for taking a player that files for arbitration all the way to a hearing abound on the team's side. Faurot (2001) elaborates on this: "arbitration hearing costs may include more than the marginal costs of preparation and attendance at a hearing. A baseball club may prefer to avoid making arguments to the arbitrator about their player's deficiencies" (Faurot 23). Therefore, only a few players will reach a hearing in a given year, rendering the sample size of players too small if I wish to maintain a sample that occurs after the influx of modern analytics. I combat this issue by comparing the markets that are comprised of the contracts of any player eligible for arbitration with those comprised of free agents. There are clear distinctions between the subsets. Free agency is effectively an open market, with each team able to evaluate a player and offer him any amount of money for any length of years. Meanwhile, a team maintains full rights to a player throughout arbitration years with the option to have a contract appropriated through the arbitration process or to negotiate with the player with the hearing as an ultimatum. Everyone involved in the process is aware of the potential hearing as an endgame and it guides negotiations, ensuring both sides prefer an agreement to a hearing before it occurs.

In both classes of players, I set a minimum at 300 plate appearances for the season prior to the contract. This has a dual purpose: to account for injury concerns that diminished player salary instead of production and to take out bench players whose statistics are compiled in few games and are not comparable to those of a starter.

As mentioned earlier, WAR is a key component of my study. WAR is included in regressions as a strong estimate of a player's value in the years preceding the contract and an example of the all-encompassing evaluation utilized by most in front offices. WAR is a component stat that tries to quantify each aspect of a player's production and compare it to the level of talent that would be freely acquirable. When looking at a hitter, it weighs outcomes of a player's plate appearances with respect to the average number of runs scored the outcome produces. Each result is placed into context for a given league and year, such that the player has an estimate of runs (not runs scored) that they contributed. This is added to measures of a player's base running and defense in terms of runs created above an average player (e.g. the percentage of times a player advances from first base to third base on a single or makes a defensive play in a difficult zone for his position to reach). Finally, the aggregate number of runs is adjusted for position (an average first baseman would produce more runs than an average catcher so the bar is higher) and scaled to be above replacement instead of above average. This ultimate sum over a full season is divided by the number of runs determined to be worth a win, giving the value that a player provides his team above what is available at minimum salary.

In essence, what is most crucial is that this formulation translates everything a player does on the field into wins. By regressing WAR and the other counting stats on a player's salary, I can approximate the residual market value attributed to these statistics beyond their impacts on wins. It is important to note that there is a relatively strong correlation between prior year WAR and average annual value (Figure 1, r=.5328). Unsurprisingly, the correlation is higher for free agents than arbitration players (Figures 2 and 3 r=.7137 vs. .431). This is another indication that a more thorough examination and accurate measure of value comes from the free agent market.

In the public sphere, free agents dominate media hype and garner the biggest contracts; these are the "established" players so impulsive thought might expect these players to be of higher quality level than their arbitration counterparts. However, a perusal of the data demonstrates that this is not necessarily the case. Looking at previous year's WAR shows that, as a whole, the arbitration eligible players were more productive, a pattern that follows suit for every other individual category as well. Figures 4-8 demonstrate this in graphical form.

The makeup of the respective classes of players helps to frame potential rationale for this finding. It has become accepted that players peak in their production levels near age 27, which means that after this age they begin to decline. Nate Silver, the noted statistician who gained his start analyzing baseball, has done extensive work on aging curves in baseball and reaffirms this conclusion, demonstrating a model that is concave down for an average career path, peaking at age twenty-seven (Silver 263).

The average debut age for major league players between 2005 and 2009 was 24.4 (Lindbergh 2). This is likely skewed with more talented players forcing their ways into the majors earlier than complimentary players, but it is rare for a player to debut prior to age 21, which would place his free agent years on track to begin at 27, on the precipice of probable decline. Further, this would only be the case for first time free agents so many signing second, third, or fourth contracts would be significantly older. Looking at the average age in each subset from my dataset help to flesh this out as free agents averaged 32.06 and arbitration eligible players averaged 27.35. Instead of including age directly in my regressions, I will use service time as a proxy. This is because the structure of contract rules gives supreme importance to differences in service time. For example, the distinction between the compensation of a player with six years of service time eligible for arbitration and a free agent with seven is enormous despite their potentially similar ages. The high correlation between the variables (.8137) situates it as an effective proxy.

Another interesting case regards players eligible for arbitration to whom a team decides not to tender a contract. These players would make up the bottom of the barrel of arbitration eligible players, as teams feel that they are not worth their potential determined salary based on prior salary level and would rather relinquish exclusive rights to remove the player from their obligations. As a result, these players end up in the free agent pool with the ability to sign with any team despite their lesser talent. They become some of the youngest and worst players in the free agent dataset, and almost all players who would expect a decrease in salary are removed from the arbitration pool. Looking at the numbers, only 6 of 289 arbitration eligible players included in the dataset actually saw a decrease in salary while 26 were not tendered a contract and became free agents (16 of whom saw decreases). While this would further contribute to the lessening of talent in the free agent database, it does not seem like it would have be harmful to results given a dummy variable for the non-tendered players. They are, after all, still subject to the conditions of free agent market so it would be misleading for them to be grouped with single-buyer arbitration-eligible players despite being more comparable in age and service time.

4. Methodology

My main strategy to compare drivers of salary in free agency and arbitration involves regressing WAR along with various counting statistics on average annual value to determine their individual impacts on salary. Average annual value (total contract value in dollars/length of contract in years) is not a perfect measure of salary as, for example, a six-year 72 million dollar contract would go to a more valuable player than a one-year 12 million dollar deal. However, it offers a fair estimate of the team's financial commitment to a player over the life of a contract. It is, for instance, far more accurate than total contract value, which would be misleading in ignoring the length of contracts, or first year salary because of the commonness of back-loaded contracts in free agency. In all regressions, I use the natural log of average annual value as the dependent variable and natural log of prior salary as an independent variable. The use of logs is to correct for the skewed right nature of the distribution of contracts as they tend to be clustered in the single digit millions, but the largest of free agent deals reach figures above twenty million dollars annually.

Before performing my regressions, I had to confront a few econometric issues that emerged from the data. By the nature of the salary structure, open competition leads to a wider range of salary values in free agency (as service time increases), which if severe enough could be lead to heteroskedasticity and invalidate my results. Figure 9 displays a scatterplot of service time and salary demonstrating this increase in range. Conducting the White test and testing the null hypothesis that there is no heteroskedasticity returns a p-value of .0028, suggesting that heteroskedasticity is a problem in my data and needs to be addressed in my regressions. However, employing robust standard errors provides a simple fix.

Another issue is the multicollinearity that exists among many of the statistics that I intended on using. This can be seen in Table 1, which includes the correlations of a number of statistics included in my initial dataset. Counting statistics in particular, such as home runs, runs scored and runs batted in, have correlations or at least .75, rendering them unusable in the same regression. To avoid this problem, I run a separate model for each of the three statistics. Other explanatory variables in each case include service time, service time squared, a dummy variable for free agents, WAR (a modern evaluative metric), games played, stolen bases, batting average, interaction terms between the free agency dummy and each of the statistics, and a dummy for non-tender cases.

To examine a slightly different angle of these results, I shift my focus to the production of players in the season following the signed contract or arbitration hearing. If free agent contracts are rewarded more heavily for estimates of production like WAR than those awarded in arbitration, it would seem to follow that these contracts would also be more predictive of WAR in future years than those of arbitration eligibles. Before examining this, I set a minimum of 300 plate appearances in the season after the deal. As with the statistics from the preceding season, most of the numbers of interest accumulate with playing time so I want to set a level high enough such that players do not have artificially low totals as a result of injury or lack of playing time. I also exclude the cases of players who were non-tendered, as they would drag down the free agent subset unjustly. I run simple regressions for ln (average annual value) on a variety of statistics, examining the relationship between the two variables. With the caveat that samples are relatively small in each case (242 arbitration seasons and 150 free agent seasons), the regressions allow for a basic indication of the effectiveness of teams and agents to award the best contracts to the most productive players.

5. Preliminary Analysis

Before running the regressions, I examine a set of graphics that attempt to isolate the effect of various common statistics on salary (specifically average annual value of contracts signed). While there are a wide set of factors that are incorporated in salary determinations, these graphs help to display whether certain measures are being more heavily compensated in either free agency or arbitration. The first step was to separate each cohort (FA vs. ARB) of players into groups of approximately equal size with a certain number of cutoffs at various production levels. For instance, I split batting average into four groups with breaks at .250, .275, and .290. An initial comparison of the mean salary with respect to batting average levels in the prior season is demonstrated in figure 10. The salaries are relatively similar throughout the groups with free agents making about twice as much in the final group. However, it is unclear the extent to which these patterns are simply an artifact of free agents being more highly compensated on the open market for their services. Figure 11 resolves this issue by creating a z-score with respect to the mean for each cohort for every observation and comparing the z-scores in each subgroup. For example, the first group for arbitration has an average salary that is .39 standard deviations below the mean for arbitration eligible players. A look at the histogram demonstrates that the increase in salary for arbitration eligible players appears linear with respect to batting average, as there is a straight increase in the average salary throughout the levels. Meanwhile, free agency decreases from groups 1 and 2 and only has a large increase between groups three and four. This seems to indicate batting average is more impactful in arbitration than in free agency as players with poor batting averages find it more difficult to garner a higher salary despite other potential attributes.

Looking at WAR in the same context, while players who produced less than 2 WAR are treated pretty similarly, free agency seems to value players above this level much more highly. Baseball-Reference.com asserts that a starter should be worth at least two wins, which unsurprisingly aligns with the change of emphasis in the free agent market. Further, the average z-score for players worth at least 3.5 wins was almost tripled in free agency with respect to arbitration. Figure 13 shows the graph and displays that the most productive free agents are receiving far and away the biggest contracts.

Looking at the z-scores for runs batted in and home runs in Figures 15 and 17 respectively show that the quintiles were treated similarly on average in each market. It is interesting to note that the large differential in the largest group that was apparent in batting average and WAR is now gone, which seem to support the conclusion that the free agent market is devaluing the measures on a relative level.

6. Results

The results from my main regression are detailed in Table 2, showing the effect of each statistic on average annual value in the free agent and arbitration markets (using an interaction term of free agents on each variable). An intuitive result is that of service time and service time squared. As a player ages, he hones his ability to perform at the highest level and after a period of gained experience, he reaches his peak. Accordingly, his compensation will increase during this period of his career. However, at a certain point, the athletic ability of the aging player begins to decrease and with it will drop his production level. Teams will expect and discover this as it occurs, and contracts they will offer thereafter shrink despite the player having more years of service. This explains the significant negative result on service time squared and creates a quadratic effect for the variable on average annual value overall.

The most important result with respect to my hypotheses shows up in the coefficients on WAR. There is a positive effect of WAR on average annual value in the arbitration case as seen by the coefficient of .047, which is significant at the 99% level. This is logical as WAR is detailing the level of production of a player, and one would expect better players to make more. However, WAR also had a large significant coefficient on free agents beyond the effect it had on arbitration eligibles (equivalent to an extra 12% increase in average annual value per each percent increase in WAR over what would be received by a similar arbitration eligible player). This indicates that these comprehensive evaluations are a better approximation of what is being considered in negotiations on the free agent market.

Moving on to the traditional statistic that I include initially, runs batted in, there is a significantly positive but small coefficient in the overall dataset, signifying that they may be slightly overcompensated beyond the player's offensive contribution through WAR. However, the interaction term between runs batted in in arbitration and free agency is not significant. Interpreting the results of the regression, a 1 percent increase in runs batted in would lead to just a 1.31 percent increase in average annual value. This contradicts the idea that runs batted in are especially looked at in arbitration, but is small enough that it seems reasonable that parties have gravitated away from giving them much credence.

Although they are highly correlated to runs batted in, I wanted to check possible effects on home runs and runs scored as well so I reran the regression replacing runs batted in with each of the other statistics; results can be seen in table 3. As was the case with runs batted in, significant results remain for the variables themselves, but not the interaction terms. A 1%

increase in home runs equates to a 2.5% increase on average annual value while a 1 % increase in runs scores leads to a 1.6% increase in average annual value.

Thinking about the small but significant increases in the traditional statistics, the coefficients could simply be the team conceding a small amount to the player for the numbers that garner widespread public acclaim while still ensuring that the majority of influence is being maintained by more reliable evaluations like WAR, especially in free agency. Another possible explanation for the lack of significance of traditional statistics in arbitration despite the purported unsophistication in the arbitrator's knowledge of baseball lies in the uniqueness of each player's background. Each party understands that it can craft a convincing argument in numerous ways and employing a variety of statistics. They are less concerned with finding the "right statistic" to convince arbitrators and instead concentrate on framing their argument around the right statistic for their case. If this were happening, no individual statistic would show up as being significantly different from the norm, but arbitrators would still be manipulated away from "fair value", which would closely correlate with WAR.

The major takeaway from the results of my study is that salaries being levied in arbitration are not as true to prior production, assuming that WAR proxies a player's worth accurately. Therefore, it would seem to follow that contracts issued in arbitration would be less predictive of future value and should show disparities from free agent contracts in terms of value seen. Results in table 4 show that, surprisingly, the larger coefficient was on the measure of production in the arbitration subset for each case. For example, in the case of WAR, β=.84 in arbitration, which translates to an increase of .008 in WAR for each one percent increase in average annual value. In free agency, β=.727, which equates to an increase of .007 for each percent increase. This is a miniscule difference, and especially given the small sample size and large fluctuation in year-to-year performance, it seems difficult to have any real confidence in making takeaways about the nature of the contracts in the markets. The other statistics have a larger difference in coefficients with arbitration being higher, but this may simply be an offshoot of the aging curve in which free agents tend to be declining from their prior numbers with arbitration players still improving. It also may be a result of the open market of free agency; since teams can only bid against each other in free agency, they may be squandering the potential value that they are gaining by using the better evaluations to outbid other teams. Therefore, while WAR would be a larger determinant of salary in free agency, it would still create situations where teams are getting more bang for their buck in terms of WAR in arbitration.

7. Conclusion

My findings carry important implications regarding negotiations that occur between players and organizations in the lead-up to arbitration hearings, as well as in free agency. The only significant variable with a large effect on average annual value was WAR in both arbitration and free agency, confirming that front offices are effectively evaluating players and devoting resources to players who have shown the capability to provide more production in the past. Further, there is a large positive interaction term between free agents and WAR, indicating that the presence of an arbitrator may be hindering the market tendency to provide even more emphasis on such comprehensive evaluations.

Meanwhile, the lack of significant interaction terms on traditional statistics with free agency implies that they may not be driving results for arbitration eligible players as much as those involved with the negotiations seems to suggest. The alternative explanation that simply providing a straightforward, easily understandable case for a player’s worth to the arbitrator seems logical, but is difficult to confirm empirically.

The better predictive power of arbitration results than free agents contracts in determining future WAR was surprising, but I have doubts of its validity as a result of the limited nature of the data. It would be interesting to see if the pattern continued in a larger sample and with a greater number of future years considered. If the results held and teams are working with fixed budgets, it might behoove them to shy away from spending as much on free agents and search for more fruitful areas to reallocate money that did not involve the obstacle of an arbitrator.

Reference List

Abrams, Roger I. 2000. *The Money Pitch Baseball Free Agency and Salary Arbitration*. Philadelphia: Temple University Press. http://site.ebrary.com/id/10395867.

Burger, John D., and Stephen J. K. Walters. 2005. "Arbitrator Bias and Self-Interest: Lessons from the Baseball Labor Market." *Journal of Labor Research* 26 (2): 267–80.

Clifton, Gregg. 2013. "2013 SABR Analytics Player Agent Panel" presented at the SABR Analytics Conference, March 9, Phoenix, AZ. http://sabr.org/latest/2013-sabr-analytics-player-agent-panel.

Coppolella, John. 2013. "Arbitration Showdown: The Insider's Intro." *Baseball Prospectus*. February 4. http://www.baseballprospectus.com/article.php?articleid=19540.

Faurot, D. J. 2001. "Equilibrium Explanation of Bargaining and Arbitration in Major League Baseball." *Journal of Sports Economics* 2 (1): 22–34. doi:10.1177/152700250100200103.

Fizel, John, Anthony C. Krautmann, and Lawrence Hadley. 2002. "Equity and Arbitration in Major League Baseball." *Managerial and Decision Economics* 23 (7): 427–35. doi:10.1002/mde.1090.

Lewis, Michael. 2004. *Moneyball: The Art of Winning an Unfair Game*. New York: W.W. Norton.

Lindbergh, Ben. 2011. "Overthinking It: Promoting Prospects." *Baseball Prospectus*. February 22. http://www.baseballprospectus.com/article.php?articleid=13018.

Marburger, Daniel R. 2004. "Arbitrator Compromise in Final Offer Arbitration: Evidence from Major League Baseball." *Economic Inquiry* 42 (1): 60–68. doi:10.1093/ei/cbh044.

Miller, P. A. 2000. "An Analysis of Final Offers Chosen in Baseball's Arbitration System: The Effect of Pre-Arbitration Negotiation on the Choice of Final Offers." *Journal of Sports Economics* 1 (1): 39–55. doi:10.1177/152700250000100104.

Miller, Phillip A. 2000. "A Theoretical and Empirical Comparison of Free Agent and Arbitration-Eligible Salaries Negotiated in Major League Baseball." *Southern Economic Journal* 67 (1): 87. doi:10.2307/1061614.

Silver, Nate. 2007. "Why Was Kevin Maas a Bust?" In *Baseball Between the Numbers: Why Everything You Know About the Game Is Wrong*, First Trade Paper Edition edition, 253–71. New York; London: Basic Books.

Appendix

Table 1: Examining Multicollinearity: Correlation of Variables

	Ln(prior sal)	serv	WAR	G	HR	RBI	SB	AVG	R	OBP
Ln(prior sal)	1									
serv	0.6001	1								
WAR	0.0821	-0.0554	1							
G	0.1922	-0.0256	0.3901	1						
HR	0.2578	0.0481	0.4473	0.4695	1					
RBI	0.2867	0.0465	0.5072	0.6565	0.8718	1				
SB	-0.0957	-0.1407	0.2963	0.2318	-0.1845	-0.0994	1			
AVG	0.0375	0.0130	0.5764	0.2536	0.1613	0.3749	0.1461	1		
R	0.2093	-0.0226	0.6687	0.7489	0.5900	0.7170	0.3797	0.4698	1	
OBP	0.1373	0.0680	0.6422	0.2325	0.3689	0.4269	0.0351	0.6955	0.5078	1
SLG	0.1434	0.0319	0.5954	0.2947	0.8352	0.7666	-0.1405	0.5434	0.5471	0.6312

Table 2:Effect on ln(Average Annual Value)
(Includes Year Fixed Effects)

VARIABLE	Coefficient	Robust Standard Error	P-Value	Corresponding % increase in Average Annual Value with 1% increase in given variable
Ln(priorsal)	.319***	.028	.000	37.57
Service Time	.098**	.040	.015	10.30
Service Time ^2	-.006***	.002	.003	-.60
FA	.039	.595	.948	3.98
WAR	.052***	.015	.000	5.34
Fa*WAR	.118***	.033	.000	12.52
Games played	.002	.002	.269	.20
FA*G	-.004*	.003	.150	-.40
RBI	.013***	.001	.000	1.31
FA*RBI	-.003	.003	.325	-.30
Stolen Bases	.000	.002	.870	0
FA*SB	.007	.005	.171	.70
Batting Average	.185	.952	.846	20.32
FA*AVG	.554	1.824	.762	74.02
Non-Tender**	-.242**	.117	.039	-21.49
N=518				
R-squared=.7212				

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Effect on ln(Average Annual Value)
(Includes Year Fixed Effects)
Same Regression as Table 1 with Runs Batted In being replaced by other traditional metrics

VARIABLE	Coefficient	Robust Standard Error	P-Value	Corresponding % increase in Average Annual Value with 1% increase in given variable
RBI	.013***	.001	.000	1.31
FA*RBI	-.003	.003	.325	-.30
WAR	.052***	.015	.000	5.34
Fa*WAR	.118***	.033	.000	12.52
R-Squared = .7212				
HR	.025***	.003	.000	2.53
FA*HR	-.000	.006	.952	0
WAR	.041**	.016	.013	4.18
Fa*WAR	.107***	.034	.002	11.29
R-Squared=.7203				
Runs Scored	.016 ***	.002	.000	1.61
FA*Runs	.002	.004	.596	.20
WAR	.035**	.016	.025	3.56
Fa*WAR	.102***	.034	.003	10.74
R=Squared=.7233				

N=518
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Examining The Correlation between Contract Size and First Year Production (Sole Independent Variable is log of Average Annual Value, Minimum Set of 300 plate appearances in season following contract)

Dependent Variable	Subset	Coefficient	Standard Error	P-Value	Corresponding Increase in Dependent Variable with 1% increase in AAV
Wins Above Replacement	Arbitration	.841***	.209	.000	.0084
Wins Above Replacement	Free Agency	.727***	.180	.000	.0072
Batting Average	Arbitration	.009***	.0028	.001	8.9e-5
Batting Average	Free Agency	.0054**	.0027	.048	5.4e-5
Runs Batted In	Arbitration	19.908***	2.383	.000	.1981
Runs Batted In	Free Agency	13.347***	2.084	.000	.1328
Home Runs	Arbitration	7.559***	.936	.000	.0752
Home Runs	Free Agency	4.186***	.837	.000	.0416
Runs Scored	Arbitration	13.775***	1.888	.000	.1371
Runs Scored	Free Agency	11.770***	1.730	.000	.1171

N=242 for Arbitration, N=150 for Free Agents
*** p<0.01, ** p<0.05, * p<0.1

Figure 1:
Scatterplot of Wins Above Replacement and Average Annual Value

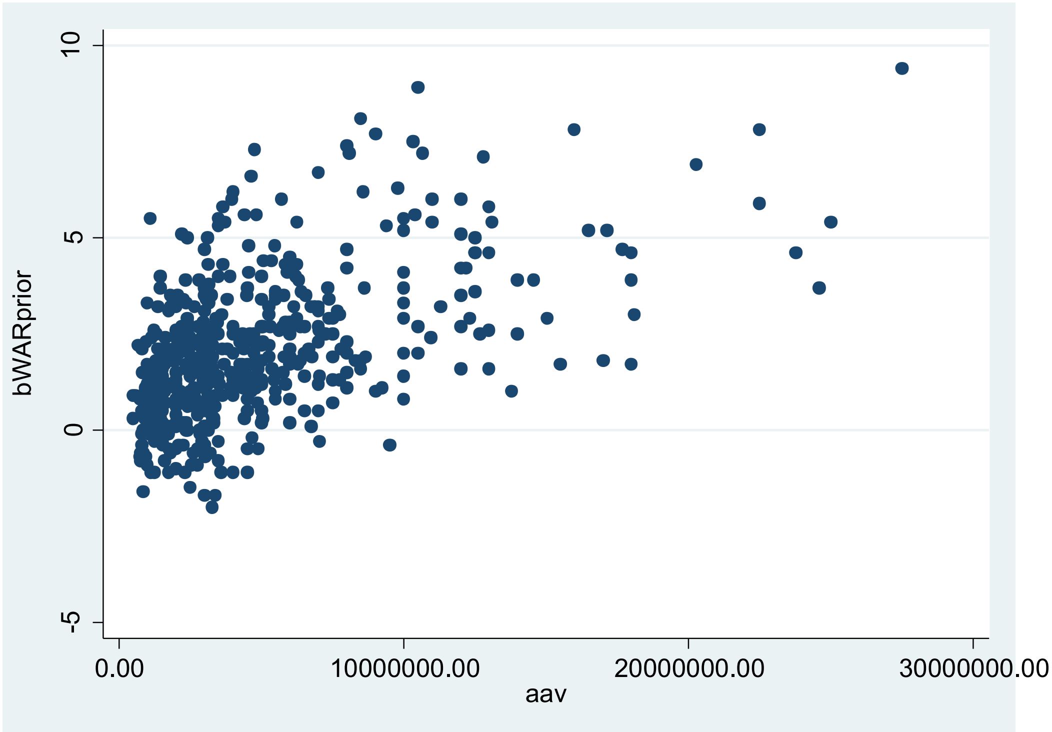


Figure 2: Scatterplot of Wins Above Replacement and Average Annual Value for Free Agents

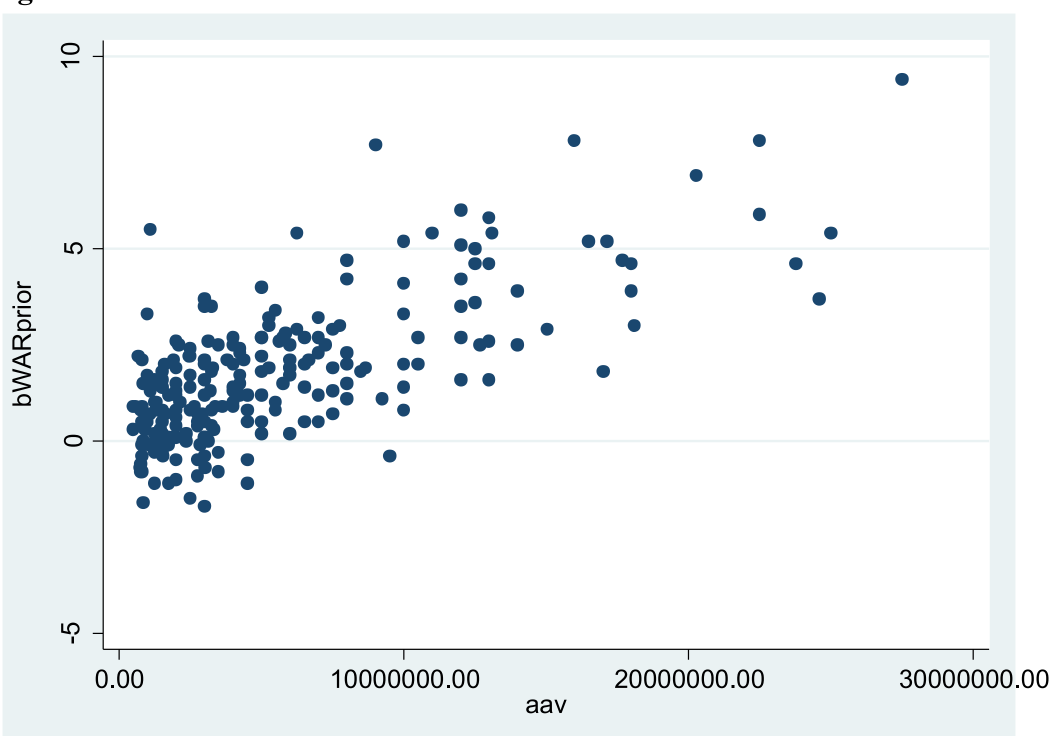


Figure 3: Scatterplot of Wins Above Replacement and Average Annual Value for Arbitration Eligibles

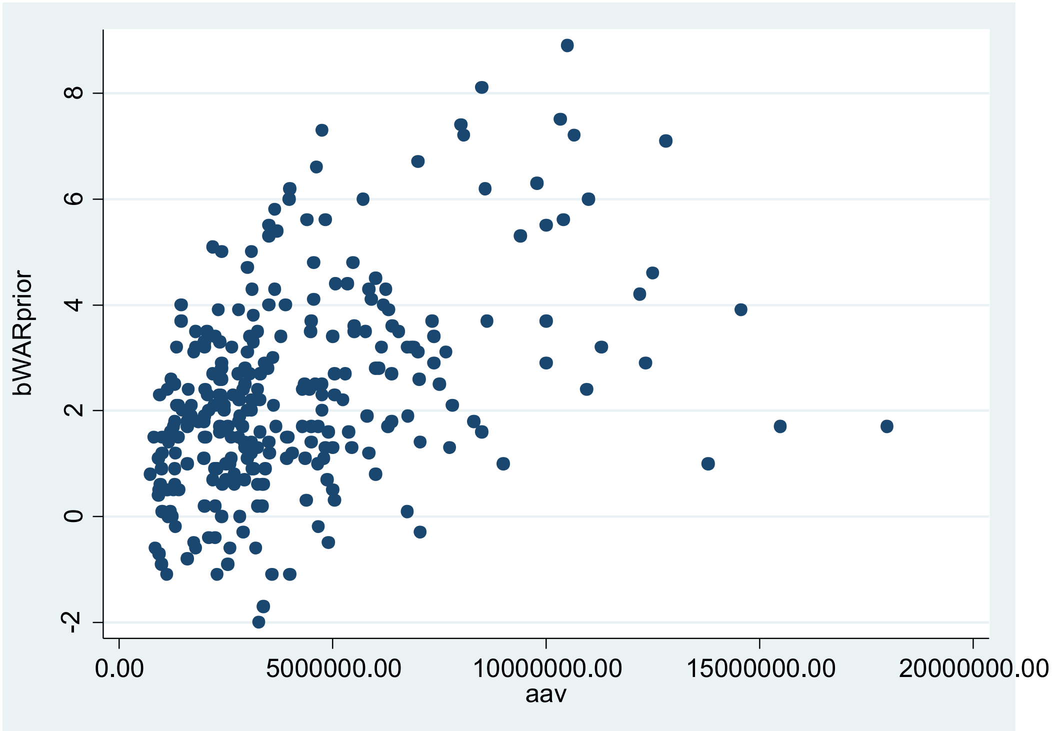


Figure 4: Box Plot of Prior Year WAR among Free Agents and Arbitration Eligibles

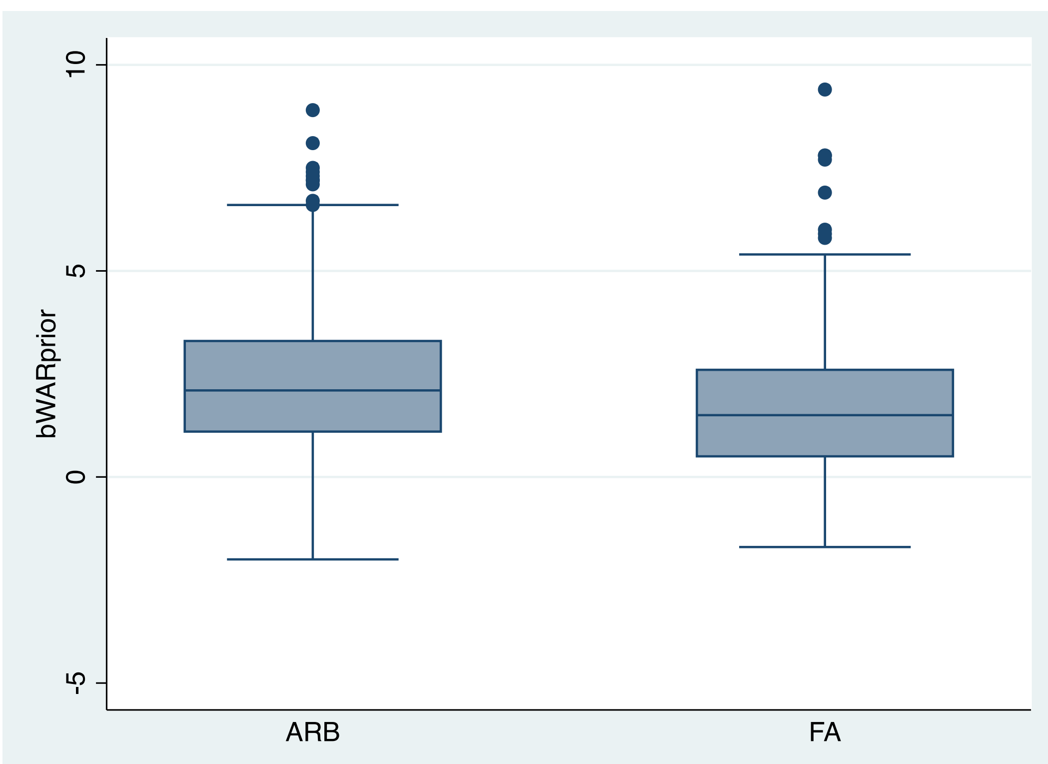


Figure 5: Box Plot of Prior Year Runs Scored among Free Agents and Arbitration Eligibles

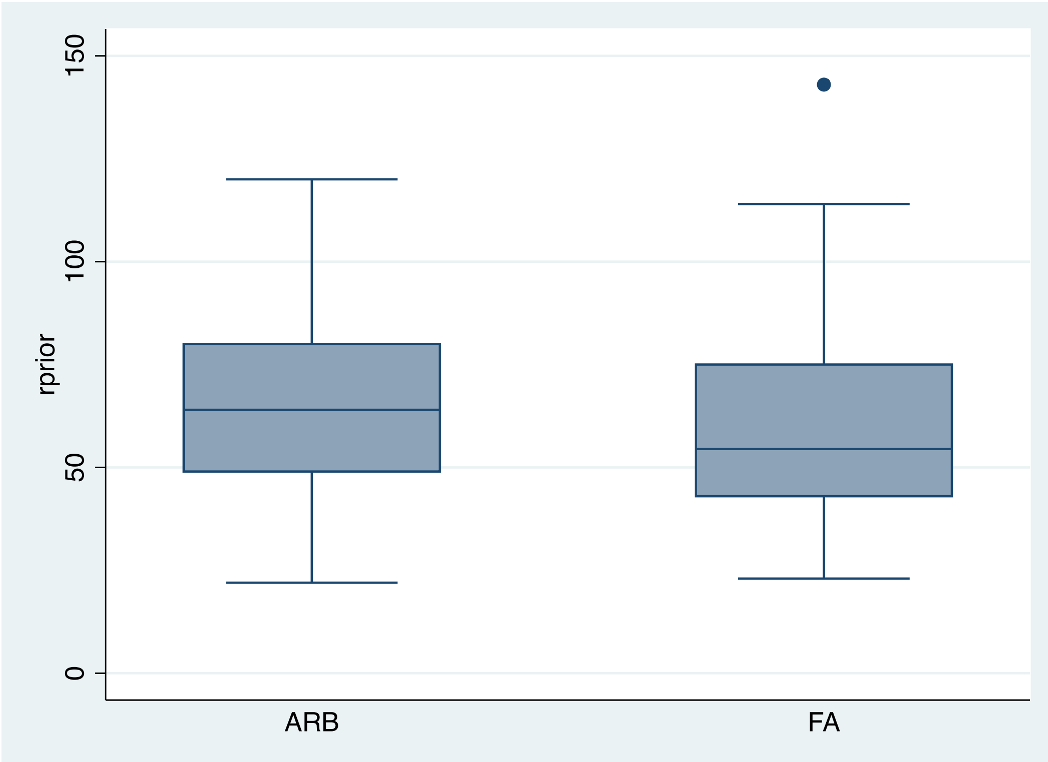


Figure 6: Box Plot of Prior Year Home Runs among Free Agents and Arbitration Eligibles

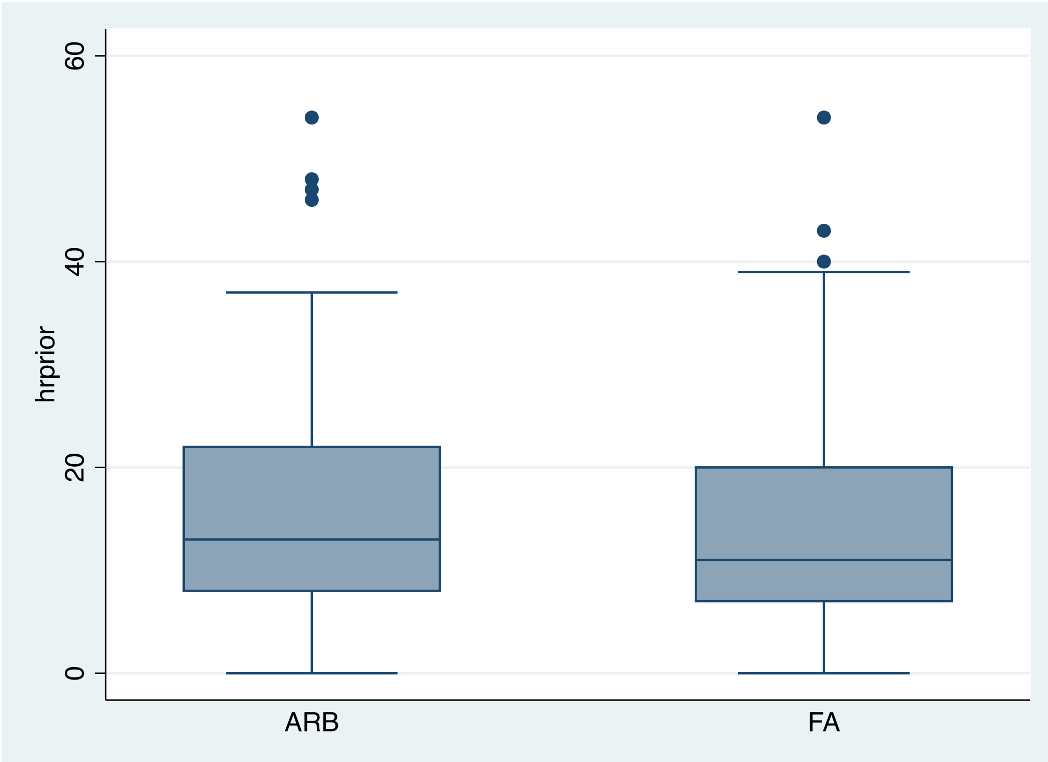


Figure 7: Box Plot of Prior Year Runs Batted In among Free Agents and Arbitration Eligibles

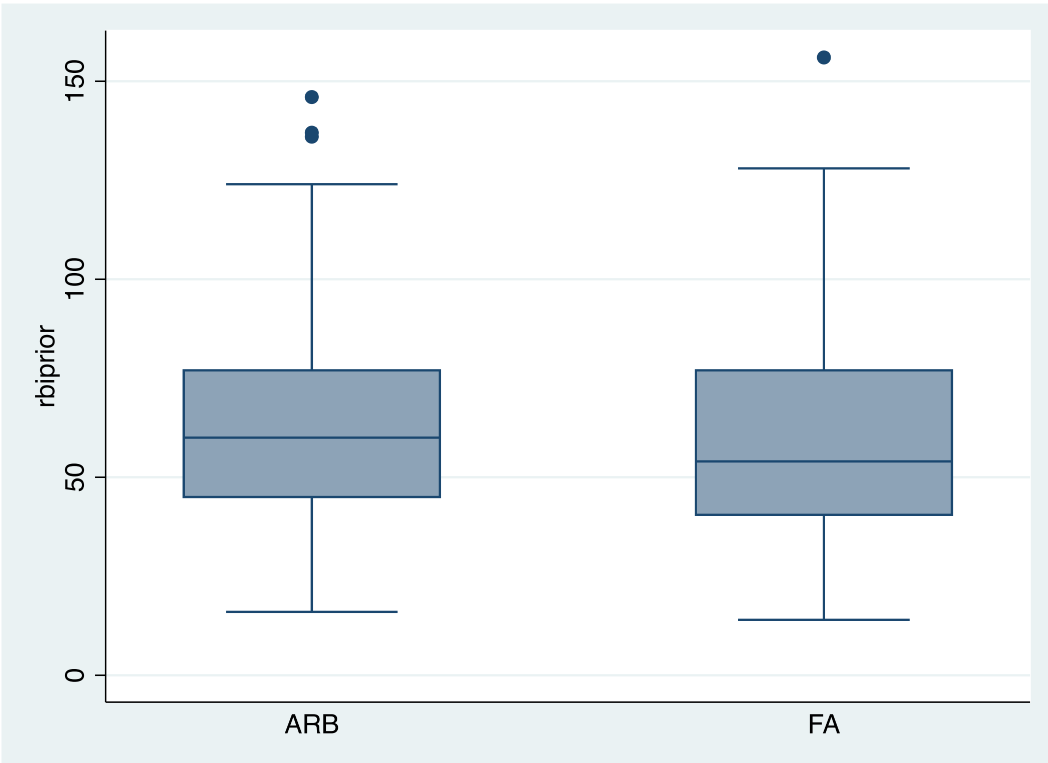


Figure 8: Box Plot of Prior Year Batting Average among Free Agents and Arbitration Eligibles

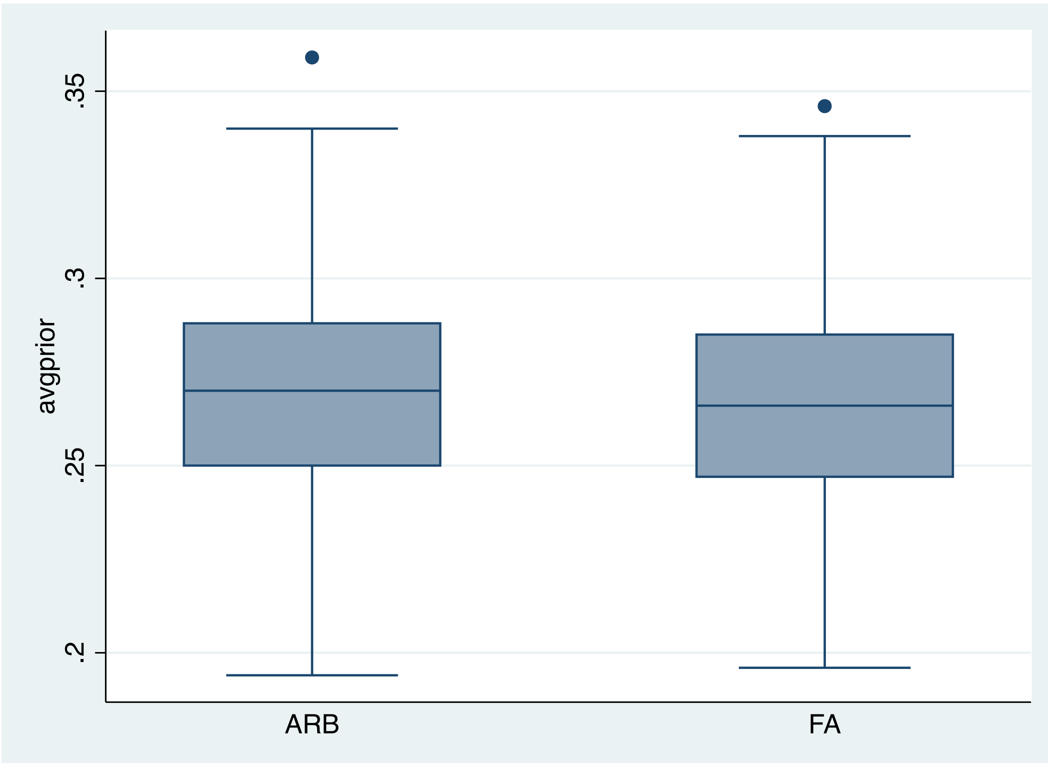


Figure 9: Scatter plot of ln(average annual value) and service time

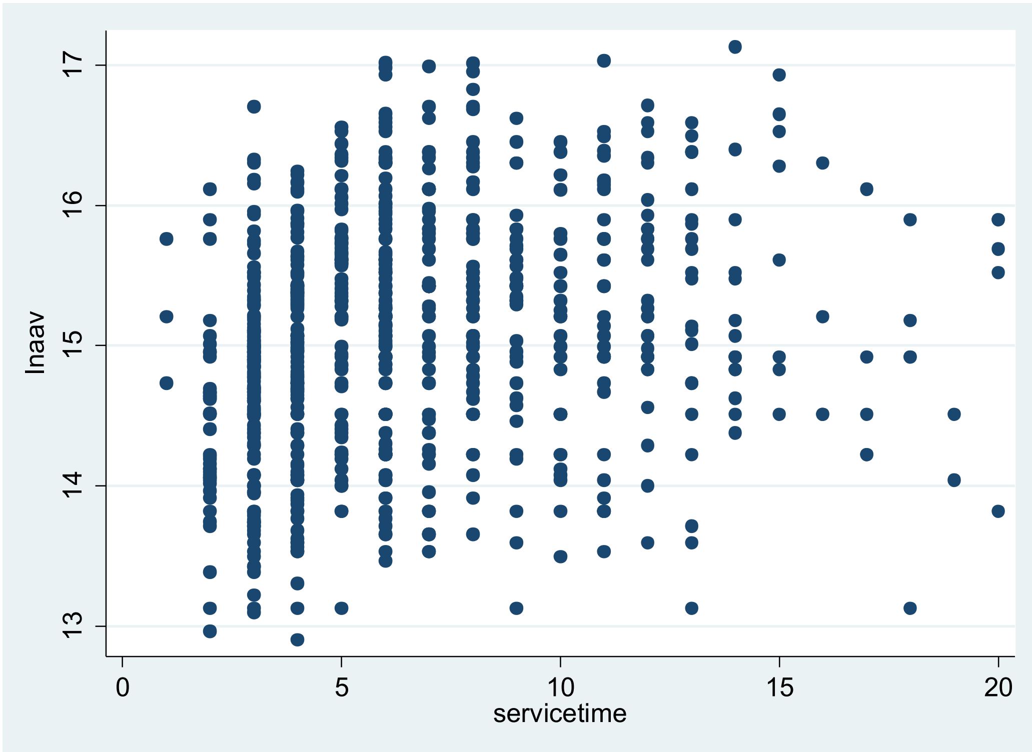


Figure 10: Comparing salaries within batting average levels in FA vs. Arbitration

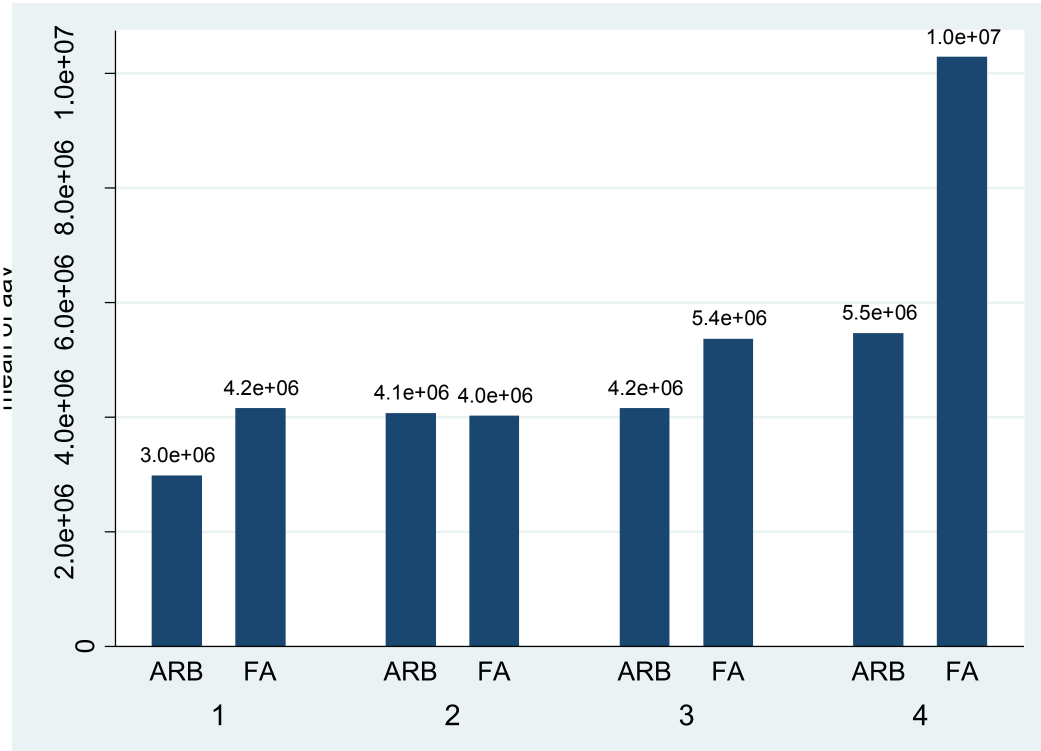


Figure 11: Comparing salaries within batting average levels in FA vs. Arbitration Using Z-Scores

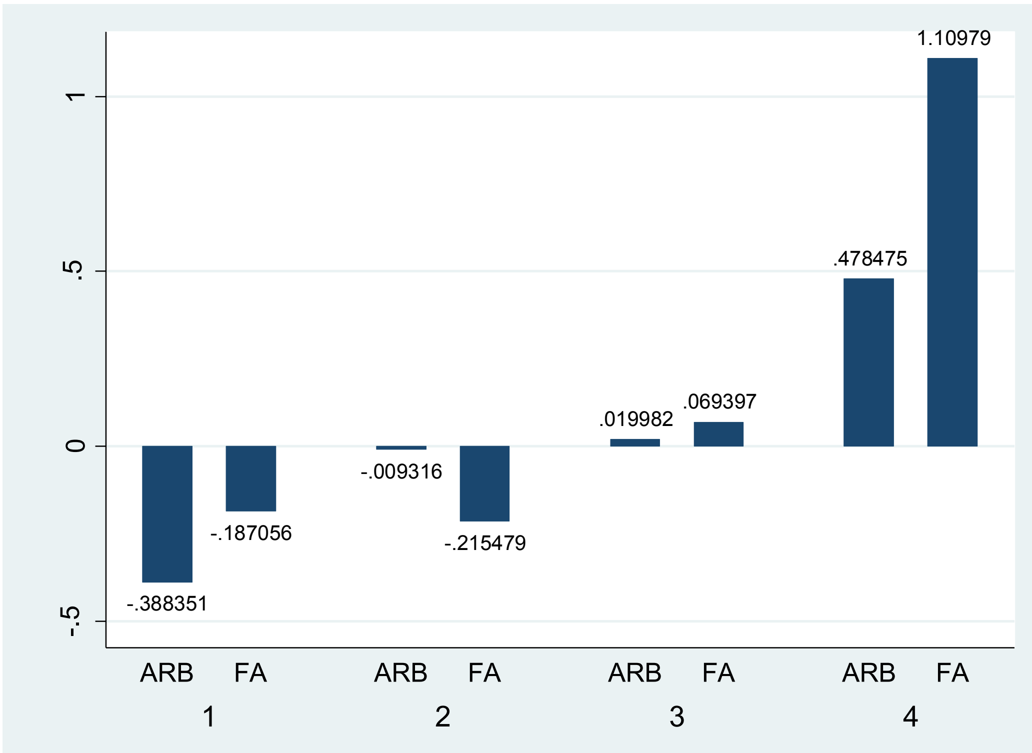


Figure 12: Comparing salaries within WAR levels in FA vs. Arbitration

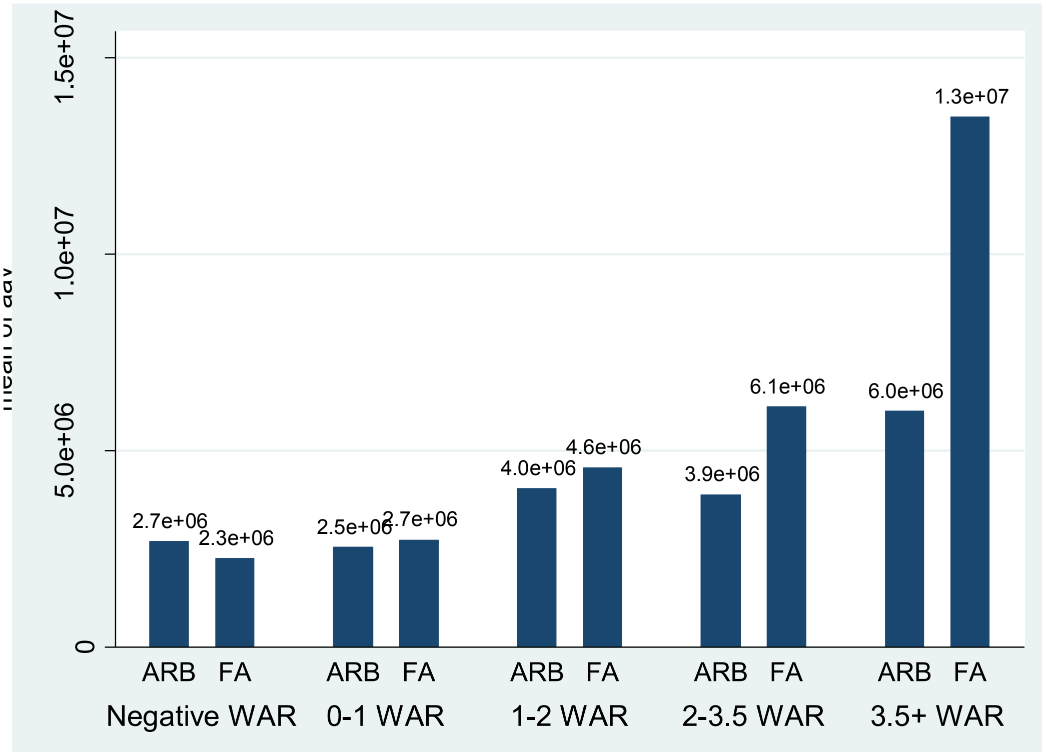


Figure 13: Comparing salaries within WAR levels in FA vs. Arbitration Using Z-Scores

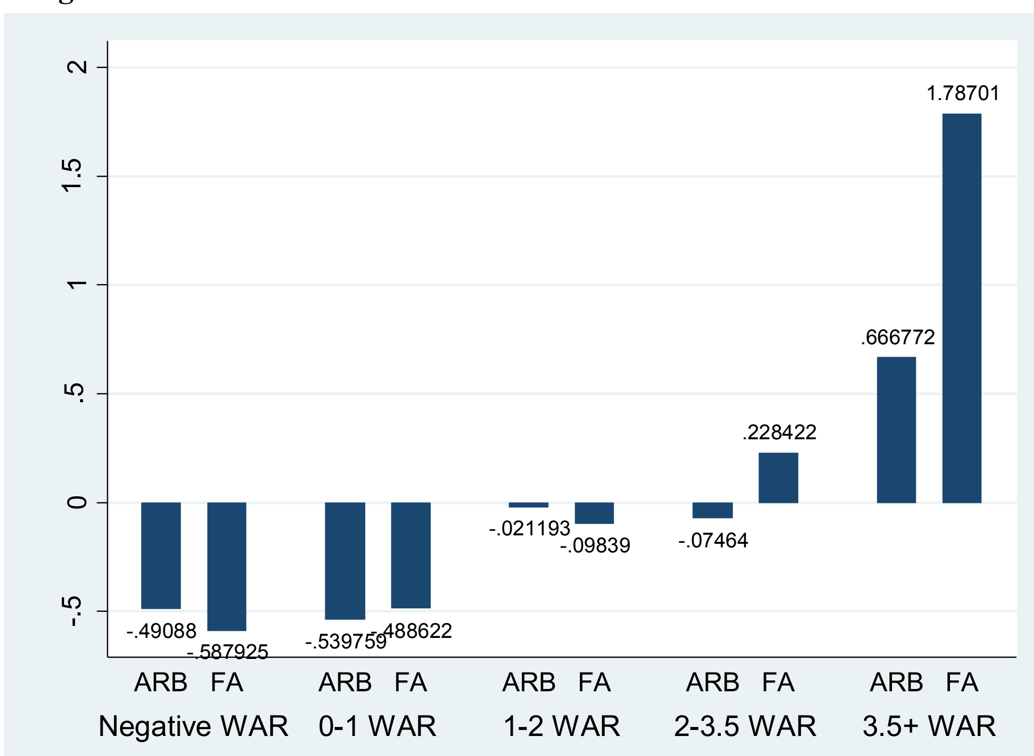


Figure 14: Comparing salaries within runs batted in levels in FA vs. Arbitration

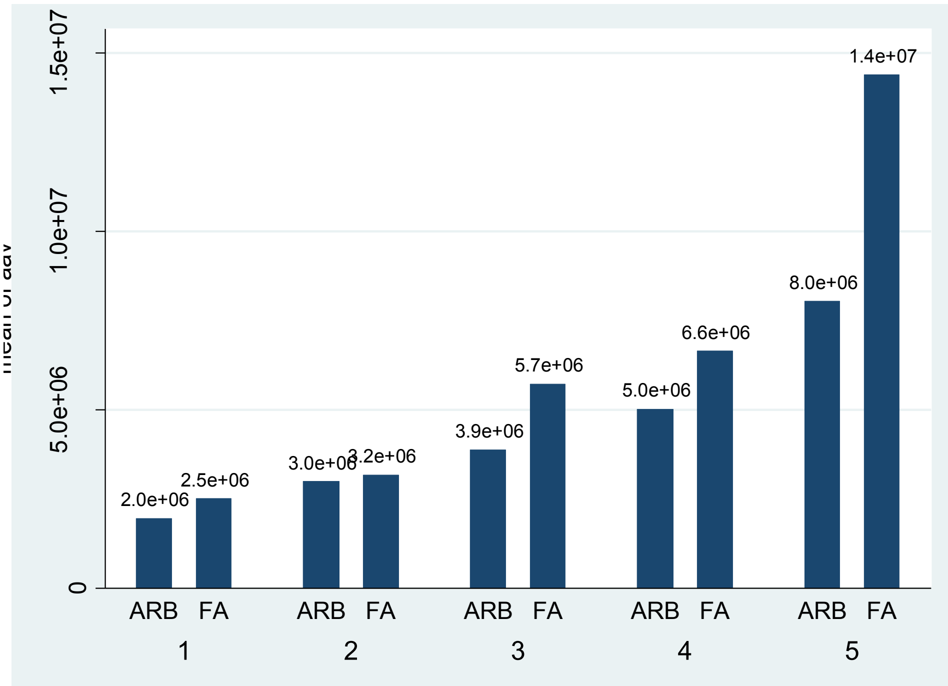


Figure 15: Comparing salaries within runs batted in levels in FA vs. Arbitration Using Z-Scores

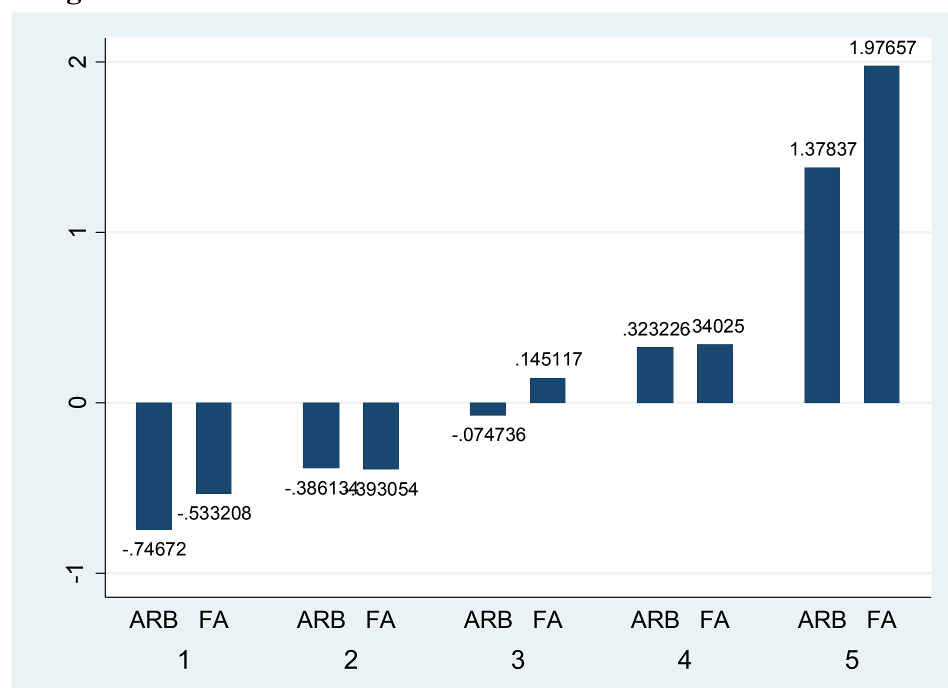


Figure 16: Comparing salaries within HR levels in FA vs. Arbitration

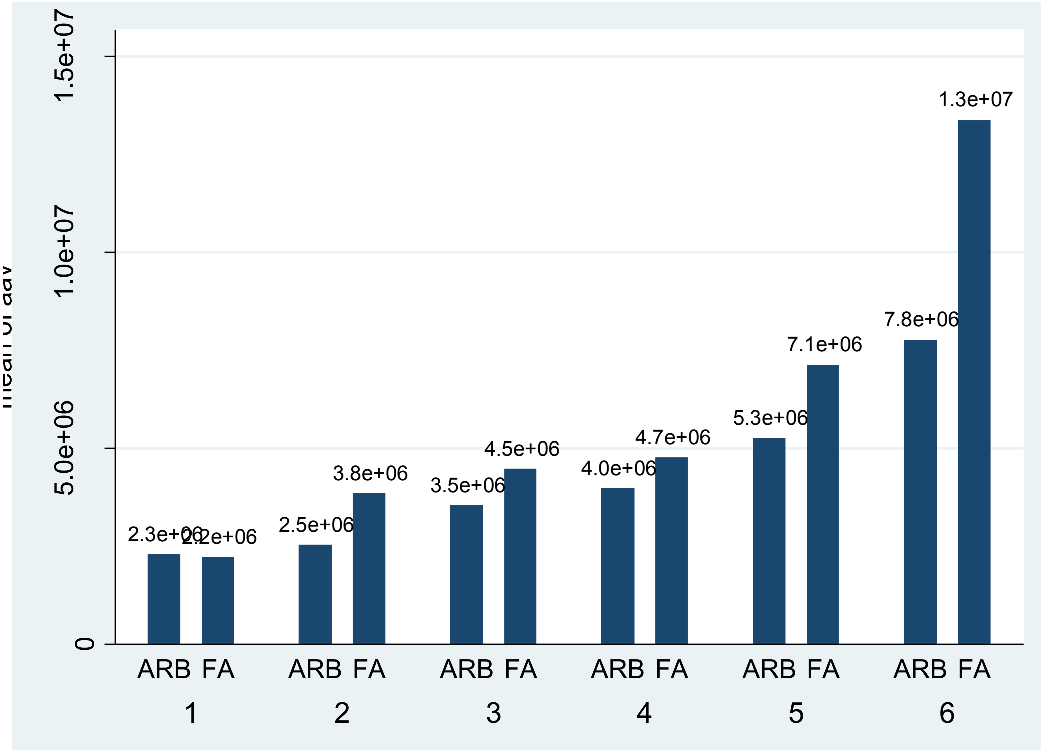
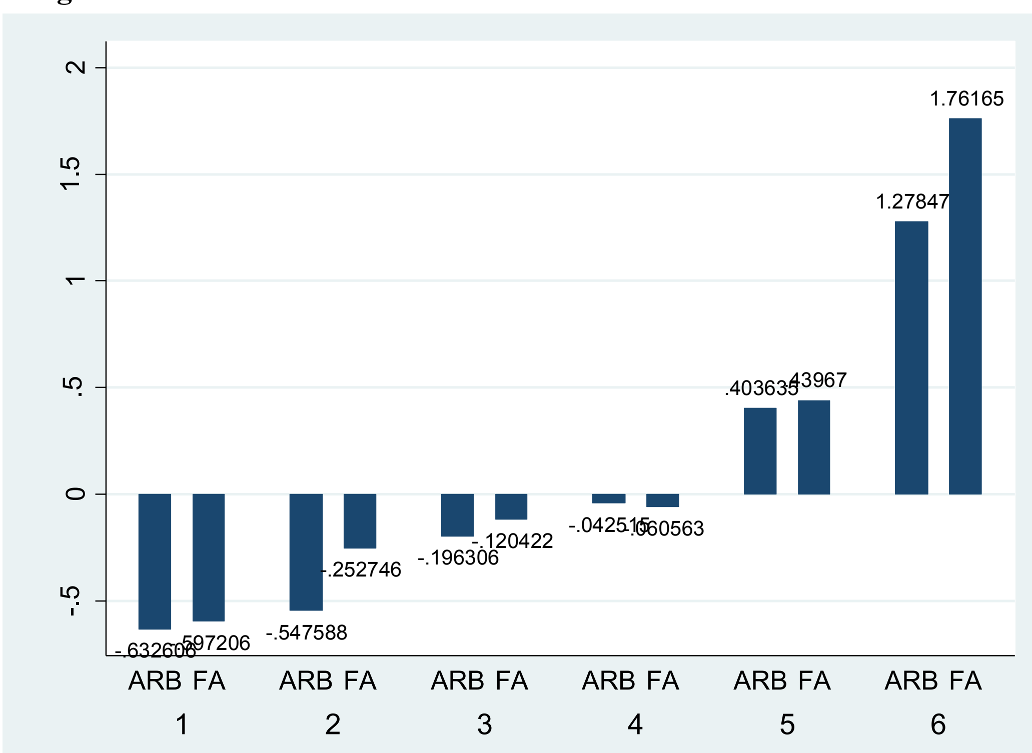
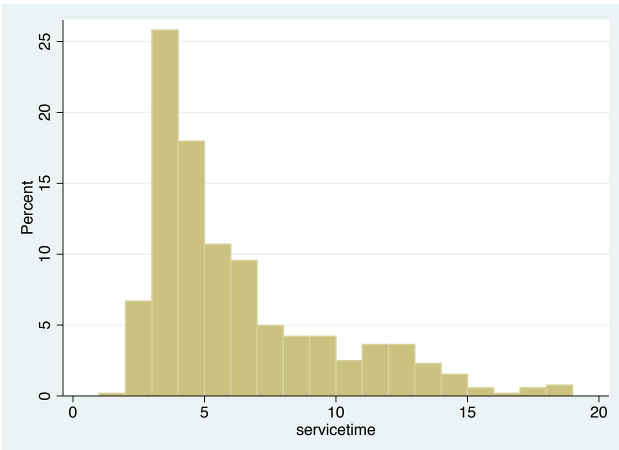


Figure 17: Comparing salaries within HR levels in FA vs. Arbitration Using Z-Scores

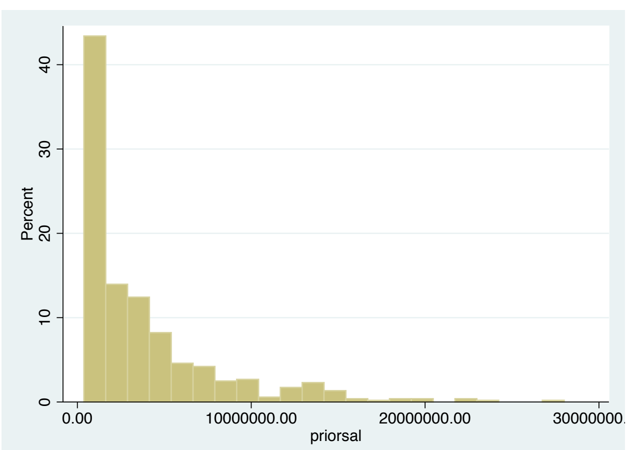


Data Appendix

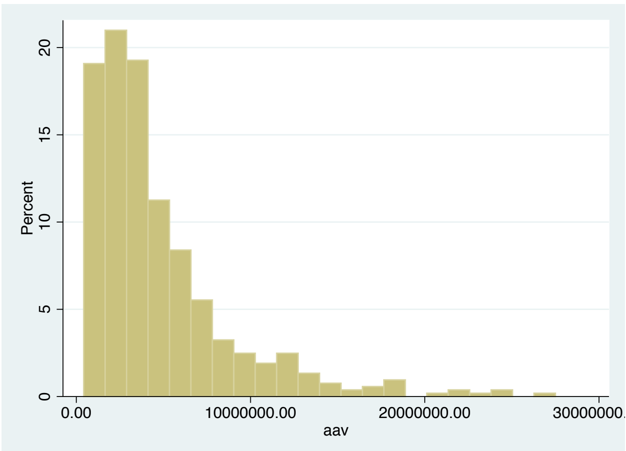
Variable Name: Service Time
Descriptive statistics:
N: 521
Mean: 5.77
Standard deviation: 3.48
Min: 1
Max: 19



Variable Name: Prior Salary
Descriptive statistics:
N: 520
Mean: $ 3,691,782
Standard deviation: $4,237,139
Min: $375,000
Max: $28,000,000

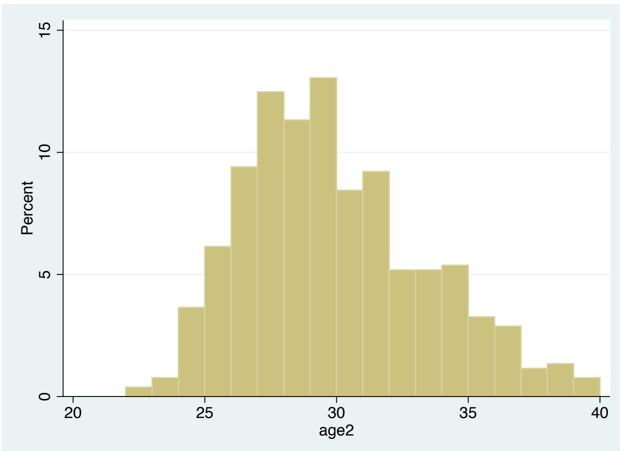


Variable Name: Average Annual Value
Descriptive statistics:
N: 520
Mean: $4,757,500
Standard deviation: $4,166,256
Min: $500,000
Max: $27,500,000

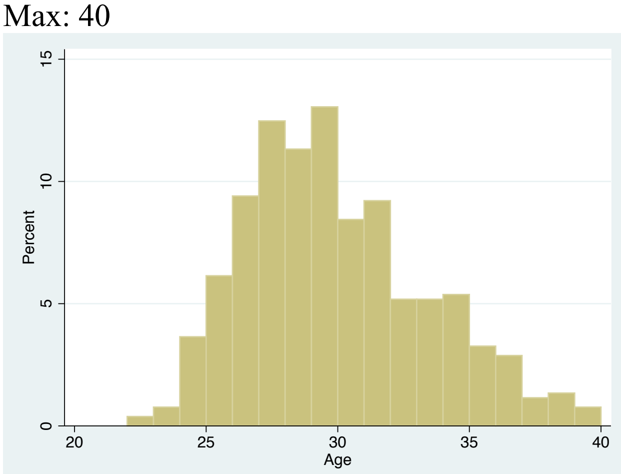


Variable Name: Year
Descriptive statistics:
N:521
Distribution:
2008: 75
2009: 75
2010: 104
2011: 95
2012: 87
2013: 85

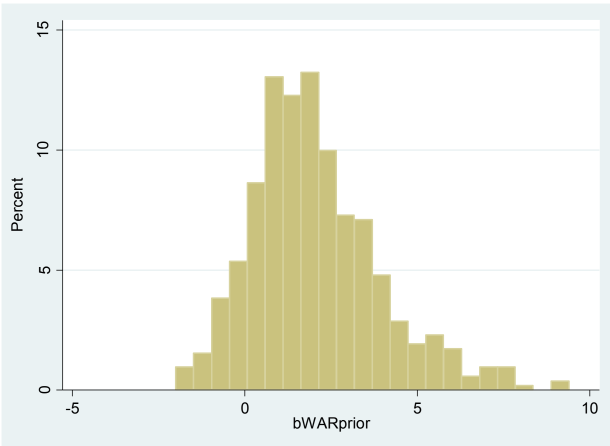
Variable Name: Age
Descriptive statistics:
N: 521
Mean: 29.44
Standard deviation: 3.46
Min: 22
Max: 40



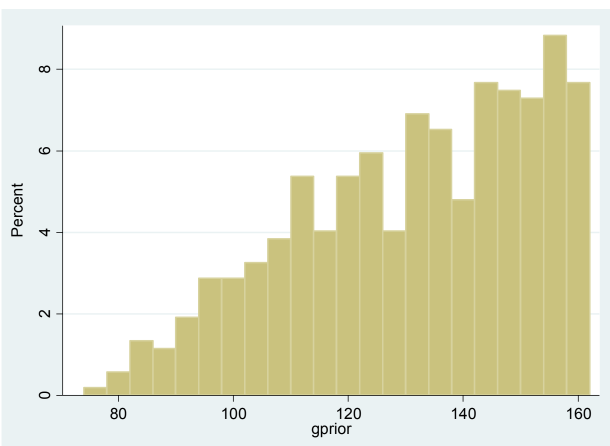
Variable Name: Age
Descriptive statistics:
N: 521
Mean: 29.44
Standard deviation: 3.46
Min: 22
Max: 40



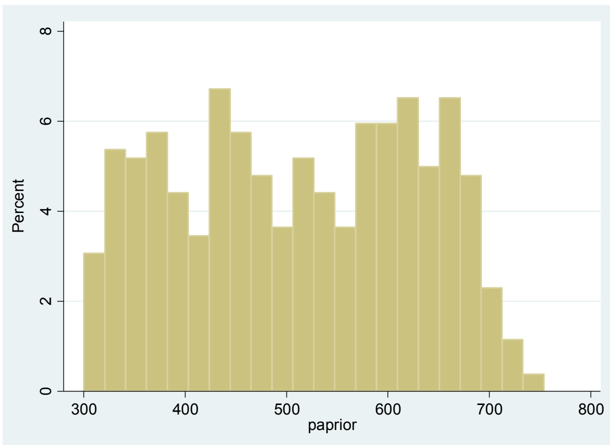
Variable Name: Wins Above Replacement – Prior Season
Descriptive statistics:
N: 521
Mean: 2.05
Standard deviation: 1.88
Min: -2
Max: 9.4



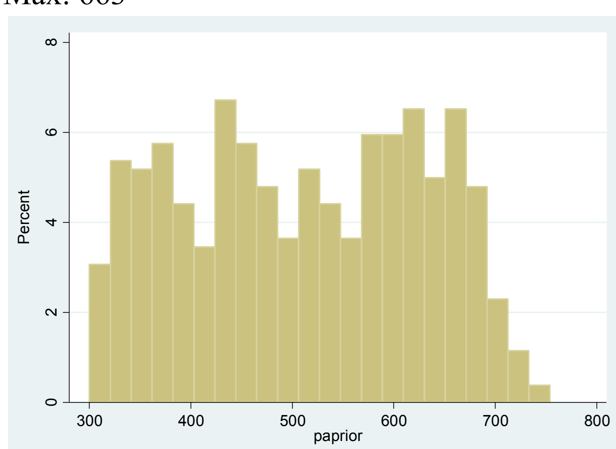
Variable Name: Games Played – Prior Season
Descriptive statistics:
N: 521
Mean: 130.96
Standard deviation: 20.94
Min: 74
Max: 162



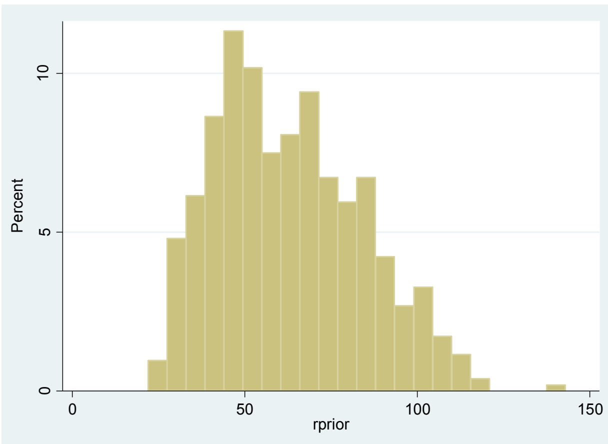
Variable Name: Plate Appearances – Prior Season
Descriptive statistics:
N: 521
Mean: 511.45
Standard deviation: 118.24
Min: 300
Max: 754



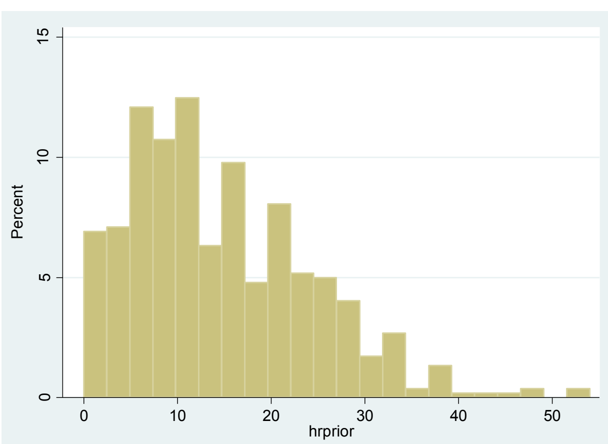
Variable Name: At Bats – Prior Season
Descriptive statistics:
N: 521
Mean: 455.82
Standard deviation: 105.71
Min: 248
Max: 663



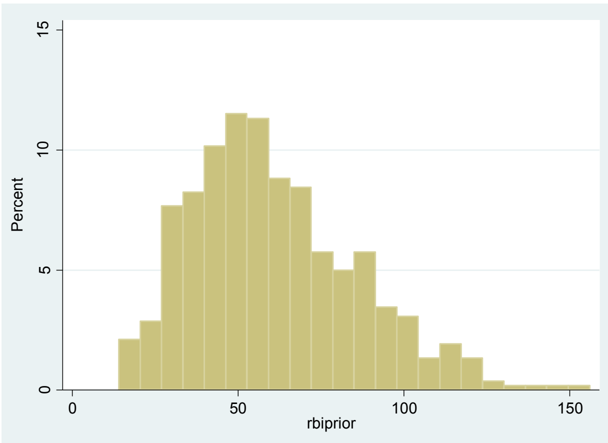
Variable Name: Runs – Prior Season
Descriptive statistics:
N: 521
Mean: 62.73
Standard deviation: 21.27
Min: 22
Max: 143



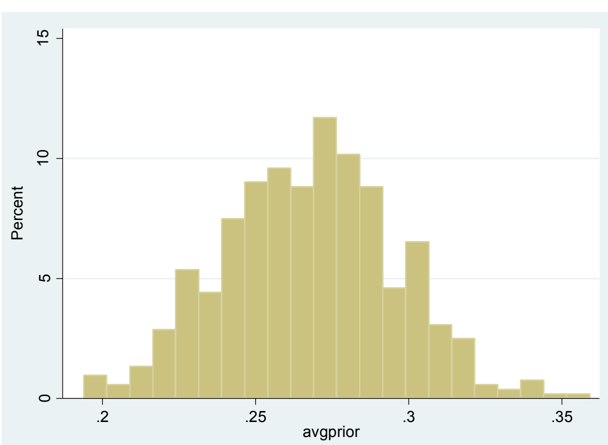
Variable Name: Home Runs – Prior Season
Descriptive statistics:
N: 521
Mean: 14.63
Standard deviation: 9.68
Min: 0
Max: 54



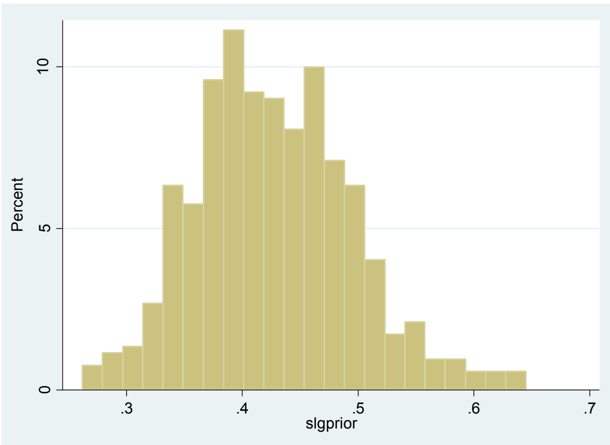
Variable Name: Runs Batted In – Prior Season
Descriptive statistics:
N: 521
Mean: 60.90
Standard deviation: 24.96
Min: 14
Max: 156



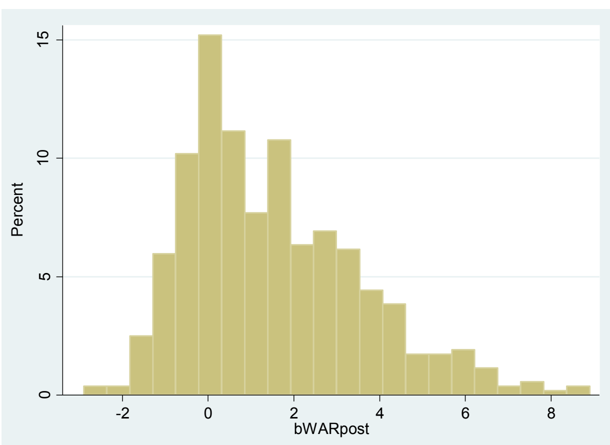
Variable Name: Batting Average – Prior Season
Descriptive statistics:
N: 521
Mean: .268
Standard deviation: .028
Min: .194
Max: .359



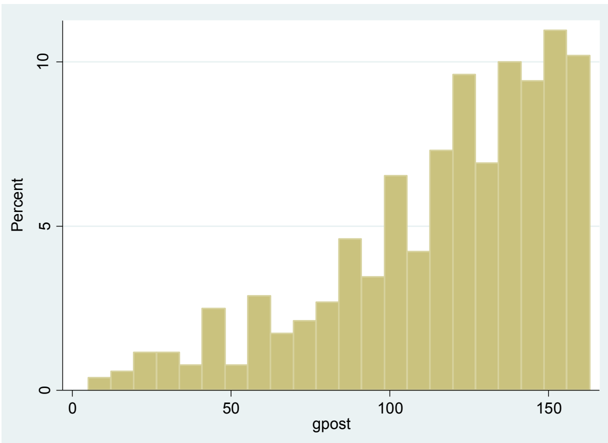
Variable Name: Slugging Percentage – Prior Season
Descriptive statistics:
N: 521
Mean: .427
Standard deviation: .068
Min: .262
Max: .645



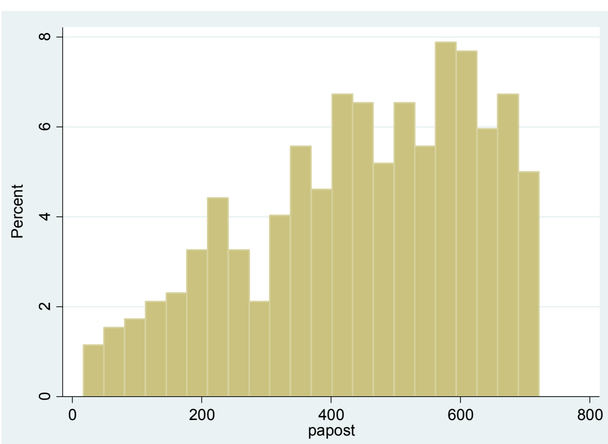
Variable Name: Wins Above Replacement – Following Season
Descriptive statistics:
N: 520
Mean: 1.527
Standard deviation: 2.032
Min: -2.9
Max: 8.9



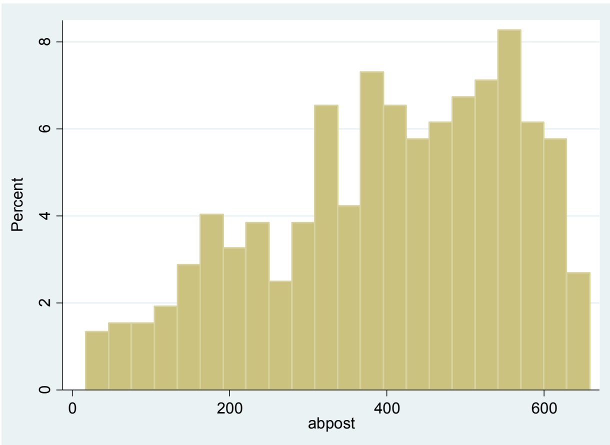
Variable Name: Games Played – Following Season
Descriptive statistics:
N: 520
Mean: 117.41
Standard deviation: 35.52
Min: 5
Max: 163



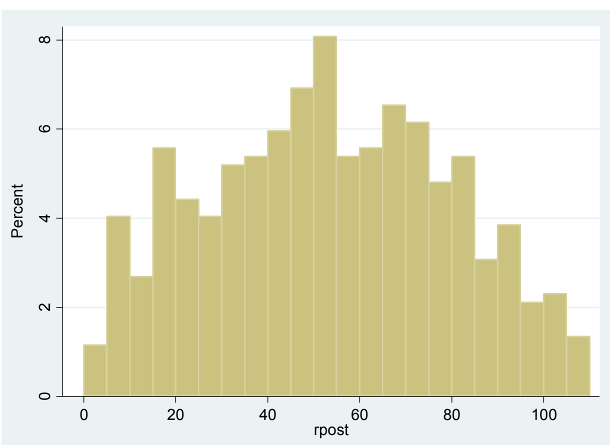
Variable Name: Plate Appearances – Following Season
Descriptive statistics:
N: 520
Mean: 448.68
Standard deviation: 175.66
Min: 17
Max: 722



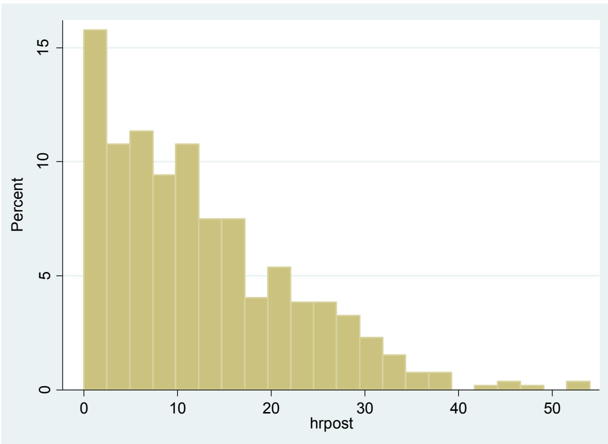
Variable Name: At Bats – Following Season
Descriptive statistics:
N: 520
Mean: 400.51
Standard deviation: 175.66
Min: 17
Max: 658



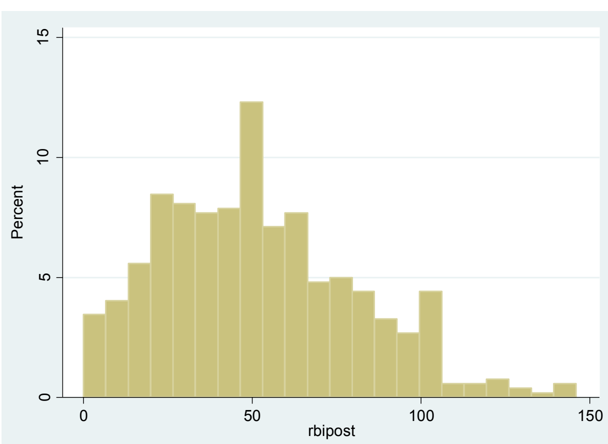
Variable Name: Runs – Following Season
Descriptive statistics:
N: 520
Mean: 52.84
Standard deviation: 26.27
Min: 0
Max: 110



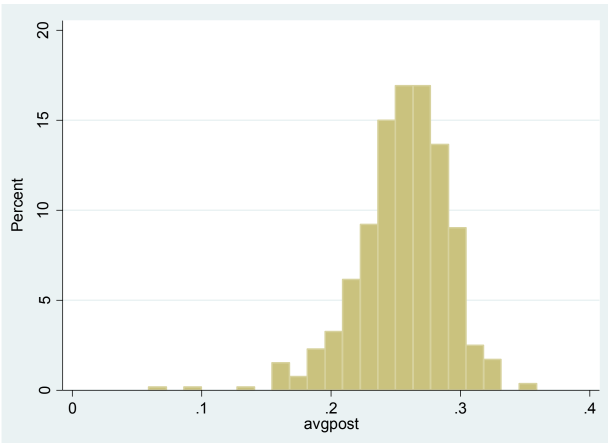
Variable Name: Home Runs – Following Season
Descriptive statistics:
N: 520
Mean: 12.38
Standard deviation: 9.98
Min: 0
Max: 54



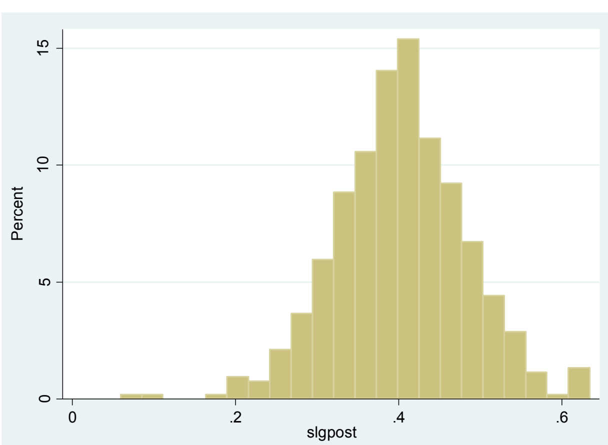
Variable Name: Runs Batted In – Following Season
Descriptive statistics:
N: 520
Mean: 51.75
Standard deviation: 28.98
Min: 0
Max: 146



Variable Name: Batting Average – Following Season
Descriptive statistics:
N: 520
Mean: .255
Standard deviation: .035
Min: .059
Max: .359



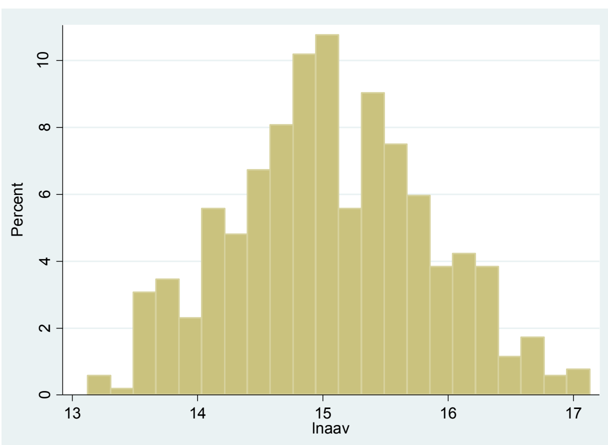
Variable Name: Slugging Percentage – Following Season
Descriptive statistics:
N: 520
Mean: .402
Standard deviation: .081
Min: .059
Max: .634



Variable: Free Agent
Dummy Variable Distribution:
0: 289
1: 232

Variable: Non-Tender
Dummy Variable Distribution:
0: 495
1: 26

Variable: Log of Average Annual Value
N: 520
Mean: 15.07
Standard deviation: .786
Min: 13.12
Max: 17.13



Variable: Log of Prior Salary
N: 520
Mean: 14.49
Standard deviation: 1.17
Min: 12.83
Max: 17.15