The Price of Winning: Quantifying Player Value in Baseball's Data Era

Liam Sit-Yee

A Thesis

In the Department

of

Finance

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science (Finance) at

Concordia University

Montréal, Québec, Canada

March 2026

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: **Liam Sit-Yee**

Entitled: **The Price of Winning: Quantifying Player Value in Baseball's Data Era**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Finance)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair
Dr. Thomas J. Walker

______________________________ Examiner
Dr. Thomas J. Walker

______________________________ Examiner
Dr. Ravi Mateti

______________________________ Supervisor
Dr. Saif Ullah

Approved by ______________________________
Dr. Frederick Davis Chair of Department

March 24, 2026 ______________________________
Dr. Anna-Marie Croteau Dean of John Molson School of Business

Abstract

The Price of Winning: Quantifying Player Value in Baseball's Data Era

Liam Sit-Yee

This research examines how Major League Baseball teams value players by integrating advanced performance metrics with detailed salary information across nine seasons (2015–2023), producing a dataset of 13,905 player-observations. Treating compensation as a financial valuation problem, the study combines Wins Above Replacement and a custom-built Win Shares dataset with contract data from Spotrac to evaluate how on-field contributions translate into monetary outcomes. The analysis applies standard econometric tools to identify the variables most closely associated with salary determination and to assess how institutional constraints, contract structures, and organizational resources shape observed compensation patterns.

The results indicate that performance metrics correlate with salary, although the strength of this relationship varies once age, signing bonuses, and career stage are considered. WAR is associated with higher compensation, yet its explanatory power declines substantially when forward-looking factors such as age and guaranteed payments are included, suggesting teams rely on expectations about future performance rather than current-season output alone. Experience emerges as the most stable predictor of salary, reflecting MLB's internal labor-market structure in which arbitration eligibility and free-agency thresholds govern compensation progression independent of measured productivity. Signing bonuses carry a consistent premium relative to annual salary, implying teams place considerable value on guaranteed, upfront compensation as a hedge against performance volatility. Team financial capacity also plays a significant role, as franchise valuation predicts individual salaries even after controlling for performance and position, indicating organizational resources influence bargaining outcomes beyond what on-field metrics alone would imply.

These patterns portray MLB's compensation system as a hybrid of market-based performance pricing and institutionally driven wage-setting. Player salaries reflect measurable contributions, career progression within a rules-based system, and the financial strength of employing organizations. By constructing a player-level Win Shares dataset and combining it to salary information, the study provides a structured foundation for examining how performance, institutional rules, and organizational finance jointly determine compensation. The findings offer insight into payroll structuring, contract design, and prospect valuation, and they highlight how financial disparities across teams shape competitive balance in professional sports.

Acknowledgement

I would like to express my deepest gratitude to my supervisor, Dr. Saif Ullah, for his continuous guidance, patience, and invaluable insight throughout the development of this thesis. His support and expertise were instrumental in shaping the direction of this research and in refining its quality at every stage.

I am also sincerely thankful to the members of my defense committee, Dr. Thomas J. Walker and Dr. Ravi Mateti, for their time, thoughtful evaluation, and constructive feedback. Their comments and suggestions have contributed meaningfully to strengthening this work.

I would like to extend my appreciation to Piper Till for introducing the idea that ultimately inspired this research topic.

I am especially grateful to my family, my mom, dad, and Ervin, for their unwavering support, encouragement, and understanding throughout this journey. Their belief in me provided constant motivation, and their patience during the demanding moments of this process made this achievement possible.

I would also like to thank my friends, San, Sabir, Damneet, Luke, Michel, Loik, Benjamin, Simon, Mohammed, Kallina, Yasmina, for bringing fun and lightness during stressful times. Their ability to help me unwind and enjoy moments along the way made this journey far more memorable.

Finally, I would like to acknowledge my extended family at Pratt & Whitney Canada, the Engineering Finance team, Camille, Élodie, Vanessa, Daniela, Mélanie, Guillaume, for their support, encouragement, and professional inspiration. Being surrounded by such a dedicated and knowledgeable group has been a meaningful source of motivation throughout my academic and professional development.

I am truly grateful to everyone who contributed, directly or indirectly, to the completion of this thesis.

God bless baseball.

Let's go Blue Jays!

Table of Contents

List of Tables

Preface

“People are overlooked for a variety of biased reasons and perceived flaws. Age, appearance, personality. Bill James and mathematics cut straight through that. Billy, of the 20,000 notable players for us to consider, I believe that there is a championship team of 25 people that we can afford, because everyone else in baseball undervalues them. Like an island of misfit toys. Billy, this is Chad Bradford. He’s a relief pitcher. He is one of the most undervalued players in baseball. His defect is that he throws funny. Nobody in the big leagues cares about him because he looks funny. This guy could be not just the best pitcher in our bullpen, but one of the most effective pitchers in all baseball. This guy should cost $3 million dollars a year. We can get him for $237,000.”

— Peter Brand to Billy Beane, *Moneyball* (2011)

Introduction

On December 11, 2023, the Los Angeles Dodgers signed Shohei Ohtani to a 10-year contract valued at $700 million, the largest player contract in professional sports to date. In the first season of this agreement, Ohtani won his first career World Series on October 30, 2024. Forty-two days later, Juan Soto surpassed this record by signing a 15-year contract valued at $765 million, which included a $75 million signing bonus. Considering only the base contractual values of these two agreements, their combined worth totals $1.465 billion, exceeding the valuations of the nine lowest-valued MLB franchises as reported by Forbes, including the Detroit Tigers, valued at $1.45 billion in 2024.

From a financial perspective, the escalation of player salaries raises a valuation question as to how long-term contracts are priced in proportion to the incremental cash flows players actually generate for their organizations. In corporate finance, capital allocation relies on disciplined frameworks that tie asset prices to expected future cash generation, yet Major League Baseball operates without comparable analytical rigor. Contract negotiations are shaped by precedent, bargaining power, and institutional rules rather than discounted cash flow logic or option pricing principles. This thesis challenges such disconnect by asking how teams implicitly value players, which performance and risk variables would belong in a rational pricing model, and how structural features of MLB's labour market such as arbitration, free agency thresholds, and revenue sharing mechanisms distort outcomes relative to competitive market predictions. The problem is ultimately one of asset valuation under binding institutional constraints, situated at the intersection of corporate finance and labour economics.

These record-breaking contracts reflect the financial dynamics in professional sports contracts. While the media often focuses on the highest-paid players, this raises important questions regarding the principles behind these agreements: are salaries accurately tied to on-field performance, or are they inflated by brand appeal, media exposure, and market size? More importantly, do these contracts accurately reflect the players' true contribution to their teams, or do they reveal asymmetries in how baseball's labour market determines a player's worth?

In a sport as steeped in tradition as baseball, player evaluations have often relied on straightforward statistics such as batting average (AVG), home runs (HR), runs batted in (RBI), and earned run average (ERA). This mirrors a broader trend across professional sports, where players who generate points or scoring opportunities tend to be valued more highly than those whose strengths are less discernible. In hockey, for example, a forward who accumulates goals and assists is often prioritized over a defensive-minded winger, despite the latter's importance to team success. Similarly, in baseball, players with strong offensive numbers are frequently favored over those who contribute primarily through defensive consistency or situational awareness. This traditional perspective began to shift with the introduction of sabermetrics, a data-driven approach to baseball analysis introduced by statistician and historian Bill James. Among his most influential contributions is the development of Win Shares (WS), a metric designed to quantify a player's total value by integrating offensive, defensive, and pitching performance into a single measure.

Although sabermetric tools like WS and Wins Above Replacement (WAR) aim to quantify a player's overall contribution, their application remains controversial. Many traditional scouts, analysts and commentators argue that such models reduce baseball to a science, overlooking the instinctive and intangible aspects that define player value. This tension between data and tradition underscores a broader divide in how performance is evaluated, with some teams openly embracing analytics and others maintaining a more conservative, experience-based approach.

In MLB, where there is no hard salary cap, wealthy teams acquire top talents due to their ability to outspend their competitors. Clubs maintaining a lower valuation must adopt more creative strategies to remain competitive. Many teams turned to sabermetrics, not necessarily as a cost-efficient solution, but because traditional approaches fail to yield the same results against better-funded opponents. However, when wealthy clubs begin to implement sabermetrics, the impact can be both immediate and significant. The Boston Red Sox serve as a prime example of how a wealthy franchise can leverage both its financial strength and analytical tools to achieve success when they ended an 86-year championship drought by winning the World Series in 2004. Under the leadership of then-newly owner John Henry, a vocal advocate of data-driven baseball, the Red Sox proved that when money and analytics align, the results can reshape the competitive landscape of the sport.

However, financial resources and team strategy alone do not fully determine competitive balance or player compensation in MLB. Player compensation is also shaped by the rules set forth in the Collective Bargaining Agreement (CBA), a negotiated contract between MLB and the Players Association that governs various aspects of the league, including labor rights, player movement, and salary structures. Among its many components, the CBA specifies how salaries are determined throughout a player's career, including league-minimum wages and salary arbitration eligibility. Players typically become eligible for arbitration after 3 to 6 years under contract, at which point they can negotiate for a higher salary based on their performance. Failure to reach an agreement between parties' results in the involvement of an independent arbitrator to decide the player's salary for the upcoming season. The most recent major agreement in 2022 maintained this framework, where a player's salary is often determined by their tenure in the league rather than their on-field performance. This explains in part why some high-performing players earn less than expected during early stages of their careers.

The introduction of modern metrics such as WAR has provided a more accurate and comprehensive method of evaluating player performance. WAR estimates how many additional wins a player adds to their team compared to a readily available replacement-level player, usually a readily available minor-league call-up or a low-cost free again who could be easily acquired. In this sense, WAR measures the incremental value a player provides beyond the baseline performance a team would expect at little cost. This metric accounts for a wide range of player attributes, including batting, fielding, and baserunning, and provides a more holistic approach to valuing a player's contribution to a team's success. Correspondingly to WS, WAR provides a standardized framework for comparing player value across positions and roles.

Nevertheless, the financial valuation of players remains a nuanced issue, with player salaries often influenced by a combination of on-field performance, marketability, and team financial considerations. While there is a clear correlation between performance and salary in some cases, discrepancies persist where players are undervalued despite their significant contributions. For instance, players who excel in less high-profile areas of the game, such as defense or on-base percentage, may not receive financial recognition as opposed to powerful superstars able to hit a fast ball high out of the ballpark. This mismatch between salary and performance presents an interesting challenge for performance analysts and front-office personnel who must navigate the complexities of compensation in the modern MLB landscape, where both roles increasingly converge in determining player value.

Interpreting these patterns involves considering how performance information enters the compensation process. Observable metrics such as WAR and Win Shares provide structured measures of contribution, but their influence is filtered through the institutional rules that govern MLB's labor market. Arbitration eligibility, service-time accumulation, and free-agency thresholds shape how closely salaries can track productivity, creating predictable departures from what a competitive market would imply. These constraints form part of the valuation environment in which teams operate and help explain why measurable performance does not always map cleanly onto observed pay.

Compensation also reflects how organizations manage uncertainty surrounding future performance. Contract duration, guaranteed payments, and the financial capacity of individual teams each capture different approaches to allocating risk. Longer commitments and larger signing bonuses convert uncertain future output into guaranteed present compensation, indicating that teams price volatility in player availability and expected performance when structuring contracts. At the same time, compensation patterns raise questions about the degree to which salaries align with rational valuation or whether systematic departures persist. Defensive specialists, players with less visible skill sets, and young prospects with uncertain development paths may be compensated in ways that diverge from their expected contributions, suggesting that structural frictions and organizational preferences continue to shape outcomes beyond what performance metrics alone would predict.

This thesis examines player compensation as a financial valuation problem shaped by the interaction of performance information, institutional rules, and organizational decision-making under uncertainty. A central consideration is how observable performance metrics such as WAR and Win Shares enter the salary-setting process and the extent to which arbitration procedures, service-time thresholds, and free-agency rules create systematic departures from what a competitive market would predict. These institutional features influence how closely compensation can track productivity and help explain why measurable contributions do not always translate directly into observed pay.

A Brief Introduction to Baseball

Baseball is a team sport played between two teams of nine players each, with the primary objective of scoring more runs than the opponent. A run is recorded when a player successfully advances around four bases, beginning at first base and returning to home plate, typically by hitting the ball and safely navigating the bases. Games are divided into 9 innings, during which teams alternate between offense and defense. The offensive team seeks to generate runs, while the defensive team aims to prevent scoring by recording outs. Outs occur when batters or baserunners fail to reach a base safely according to the rules of play, such as when a fielder catches a batted ball before it touches the ground, tags a runner with the ball while they are not on a base, or forces a runner out by touching a base ahead of them with the ball in hand.

Offensive performance in baseball captures how players generate runs, primarily through hitting and base running. Traditional hitting statistics such as batting average (AVG), measuring the proportion of at bats (AB) resulting in hits (H), runs batted in (RBI), quantifying the number of runs a batter drives in, and home runs (HR), occurring when a batter circles all bases on a single play, reflect core offensive outcomes but provide only a partial view of overall contribution. Base running further enhances offensive value by allowing players to advance into scoring positions and convert opportunities into runs. To address the limitations of conventional metrics, advanced sabermetric measures such as Runs Created (RC) combine a broader set of offensive events, including hits, walks, and stolen bases, into a single estimate of a player's total impact on run production. At the team level, offensive output is also shaped by strategic decisions including sacrifice attempts, hit and run plays, and base stealing, all of which influence both the frequency and efficiency of scoring opportunities.

Defensive fielding performance in baseball captures how teams prevent runs through the effective execution of fielding responsibilities. Each defensive player occupies a designated position, including the infield roles of first base, second base, third base, and shortstop, the catcher behind home plate, and the outfield positions of left, center, and right field. Modern evaluation of fielding relies on advanced metrics such as Defensive Runs Saved (DRS), quantifying the number of runs a player prevents relative to an average defender at the same position, and Ultimate Zone Rating (UZR), assessing how effectively a fielder covers their assigned area by incorporating putouts (PO), assists (A), errors (E), and the difficulty of attempted throws. These measures provide a more comprehensive understanding of defensive value than traditional statistics, offering insight into both individual fielding skill and overall run prevention.

Pitching performance in baseball represents a central component of run prevention, as pitchers work to generate outs and limit opposing offenses. Standard pitching statistics include innings pitched (IP), measuring the number of innings a pitcher completes, and earned run average (ERA), the average number of earned runs a pitcher allows per nine innings excluding the plays resulting from defensive errors. ERA isolates the pitcher's contribution to run prevention by attributing responsibility only for outcomes deemed within the pitcher's control. These measures provide a consistent basis for comparing pitchers across teams and seasons. Effective pitching also relies on strategic variation in pitch type, velocity, and location to exploit batter tendencies, and managers frequently adjust pitching personnel to optimize matchups in high-leverage situations.

Game outcomes in baseball emerge from the interaction between individual skill, situational context, and strategic decision-making. A player's contribution is rarely isolated; rather, it is shaped by the game state, including baserunner pressure, defensive alignment, and managerial choices. Hitters face different expectations depending on whether runners are in scoring position, while pitchers adjust their approach based on batter tendencies and leverage situations. Teams further refine their defensive strategies through data-driven positioning, aiming to convert batted balls into outs more efficiently. These layers of context illustrate that performance in baseball is both dynamic and interdependent, influenced by far more than isolated statistical events.

Evaluating this complexity requires a combination of traditional statistics and modern analytical tools. Counting measures capture the frequency of key events, while rate statistics assess how effectively players convert opportunities into productive outcomes. More advanced metrics, including WAR and WS, synthesize offensive, defensive, and pitching contributions into unified estimates of player value. By integrating multiple dimensions of performance, these measures offer a more comprehensive understanding of how players influence team success. This multifaceted approach to evaluation provides the foundation for the empirical analysis that follows, where performance metrics are examined in relation to salary outcomes across MLB.

Literature Review

Research on compensation and productivity in labor economics provides several frameworks for understanding how performance relates to pay. Incentive theory outlines how performance-based contracts, rank-order rewards, and observable output shape effort and compensation outcomes (Lazear, 1981; Lazear & Rosen, 1981; Prendergast, 1999). Empirical work identifies conditions under which compensation responds only weakly to productivity, as institutional constraints, risk considerations, and measurement limits restrict the sensitivity of pay to output (Medoff & Abraham, 1980; Jensen & Murphy, 1990; Oyer & Schaefer, 2011). Superstar models add a further dimension by explaining how scalable performance and concentrated demand generate steep earnings gradients from relatively small differences in talent (Rosen, 1981; Rosen & Sanderson, 2001). Together, these perspectives outline the economic forces that govern compensation systems in environments where output is both highly visible and economically consequential.

MLB incorporates these mechanisms within a single labor market as research document the influence of monopsony power, restricted mobility, revenue sharing, arbitration rules, positional differences, and evolving valuation methods on salary formation (Scully, 1974; Kahn, 1993; Fort & Quirk, 1995; Krautmann, Gustafson & Hadley, 2003; Brown, 2009; Hakes & Sauer, 2006). Such institutional features interact with measurable performance in ways shaping bargaining leverage, contract duration, and the distribution of earnings across players. More recent work integrates detailed performance metrics, contract structures, and team-level financial conditions into unified empirical frameworks for analyzing salary variation (Berri, Leeds & von Allmen, 2015; Simmons, 2022; Butler, Farnell & Simmons, 2024). With performance measured at an exceptional level of precision and compensation conditioned by league-specific rules, MLB provides a setting well suited for assessing how metrics such as WAR and WS correspond to salary outcomes.

Lazear (1981) contribution to agency and incentive theory establishes a formal foundation for understanding how performance-based compensation can align individual effort with organizational objectives. When compensation increases the marginal return to productive activity and output is observable, workers have stronger incentives to supply higher levels of performance, leading to measurable gains in productivity. Firms can also limit shirking by designing wage profiles in which early career workers are paid below their marginal product and later rewarded above it, creating long term incentives that sustain effort over the employment relationship. Lazear and Rosen (1981) introduce tournament theory to show that compensation systems based on relative performance create steep reward differentials that magnify effort incentives even when absolute output is difficult to contract on. The structure of tournament rewards means that small differences in performance can translate into large differences in pay when compensation is tied to rank, a feature that induces strong competitive behavior among workers and sharpens the link between effort and expected compensation. Prendergast (1999) develops a related perspective by examining how organizations structure incentive contracts to align worker behavior with measurable output, particularly in environments where performance differences carry meaningful economic consequences. MLB reflects these dynamics clearly: players in their initial service years earn salaries far below their on-field contributions despite often producing substantial surplus value, while underperforming players experience reduced playing time as managers reallocate opportunities toward more productive teammates. The displacement of veteran players whose productivity falls below their wage similarly mirrors the employment adjustments predicted by incentive theory. Professional sports therefore create a setting in which the central features of Lazear's framework such as observable output, incentive driven compensation, and employment responses tied to productivity are especially transparent. This incentive-based foundation offers a rigorous basis for assessing how metrics such as WAR and WS translate into compensation outcomes in professional baseball.

Performance-based compensation does not always generate the strong incentives predicted by theory, a point underscored by evidence showing that pay often responds only weakly to measurable changes in productivity. Medoff and Abraham (1980) use personnel records and performance evaluations from large firms and show that wage growth is driven primarily by seniority rather than measured output. The results point toward internal labor markets in which tenure plays a dominant role in shaping earnings, limiting the strength of performance-based incentives and reducing the extent to which compensation motivates additional effort. Jensen and Murphy (1990) study executive compensation contracts and document a similarly weak relationship between pay and performance at the top of the corporate hierarchy. Even substantial changes in firm value translate into relatively small adjustments in managerial compensation. The evidence highlights structural features of executive pay, including risk considerations, governance constraints, and the difficulty of writing complete contracts, all of which contribute to muted pay responses despite publicly observable performance and large financial stakes. Oyer and Schaefer (2011) draw on a wide range of empirical settings and identify consistent patterns in which incentive systems generate limited behavioral responses. Across diverse labor markets, the evidence points toward measurement challenges, multitasking concerns, risk aversion, and institutional constraints that weaken the link between performance and compensation. Even when output is measurable and rewards are formally tied to performance, the observed incentive effects are far weaker than theoretical models predict. A comparable tension appears in professional sports, where salaries are expected to reflect on field output but are also shaped by institutional rules, bargaining positions, and market conditions that can limit the responsiveness of pay to performance. MLB illustrates this dynamic clearly, as service time regulations, arbitration procedures, and long-term contract structures can restrict how closely compensation tracks individual productivity even when output is measured with exceptional precision. Although the Jensen and Murphy analysis predates modern sabermetric measures such as WAR and the granular tracking data introduced through Statcast, the framework outlines conditions under which compensation may diverge from observable productivity. This perspective positions MLB as a setting where the strength of the link between performance and pay can be examined directly through measures such as WAR and WS.

Rosen (1981) theory of superstars establishes a central foundation for understanding compensation patterns in MLB, as it formalizes how markets with scalable performance and concentrated demand generate extreme wage dispersion even when underlying differences in ability are modest. The framework shows that when consumers place high value on quality and performers are imperfect substitutes, small increments in talent can produce disproportionately large economic returns, creating earnings distributions that rise convexly with ability. Examples from musicians, actors, and elite athletes illustrate how performance can scale to large audiences at minimal cost, allowing top performers to capture a substantial share of total market value. Rosen and Sanderson (2001) apply these ideas directly to professional sports and document how scalable performance, media exposure, and concentrated consumer demand magnify earnings at the top of the talent distribution. MLB reflects these structural conditions, since elite players generate competitive value on the field while simultaneously driving commercial revenue through attendance, television exposure, and sponsorship demand. The resulting right skewed salary distributions follow naturally from the economics of scalable performance and concentrated demand, creating a setting in which the alignment between compensation and measurable productivity can be examined directly. In an environment where individual output can be quantified with exceptional precision, the superstar model suggests that deviations between performance and pay are not solely the product of institutional constraints but may reflect fundamental features of talent markets, making MLB an ideal context for empirical evaluation.

The first systematic attempt to quantify player value in professional sports was developed in Scully (1974) analysis of MLB salaries, which applied a marginal productivity framework to estimate each player's economic contribution. The model specified team performance as a function of individual output and converted that performance into team revenue, allowing player value to be expressed as marginal revenue product and establishing a competitive benchmark for salary determination. This structure formalized the conclusion that compensation in a competitive labor market should equal the revenue generated by a player's marginal contribution to team success. Scully's central hypothesis presented the MLB labor market as a monopsony in which players were bound to their clubs through the reserve clause and lacked any meaningful outside options. Consistent with this hypothesis, estimates generated under the model indicated that actual compensation during the reserve clause era fell substantially below competitive levels, revealing systematic underpayment caused by the absence of player mobility. Because players were contractually bound to their teams and unable to negotiate with alternative employers, clubs operated in a labor market where they held all bargaining power and captured most of the surplus generated by player performance.

Kahn (1993) identified clear shifts in salary determination following the introduction of free agency and long-term contracts. Arbitration and free agency eligibility were associated with higher salaries, while free agency also correlated with longer contract durations. The pattern aligned with Nash bargaining predictions in which longer contracts appear when teams face a credible risk of losing players to the open market. The estimates further indicated a winner's curse pattern, as the longest contract extensions appeared in situations where teams held minimal contractual control and faced the greatest risk of losing players to free agency. This concentration of long-term deals under weak bargaining conditions reinforced the role of leverage, rather than performance alone, in shaping contract duration.

Hall, Szymanski, and Zimbalist (2002) provide the empirical foundation for assessing the dynamic relationship between payroll and team performance through Granger causality tests applied to MLB data from 1980 to 2000. Analysis of this relationship indicates no causal link in either direction over the full sample period, as variations in spending did not systematically precede changes in on field outcomes, nor did past performance reliably shape subsequent payroll decisions. The correlation between payroll and performance strengthened during the 1990s, but the underlying direction of influence remained indeterminate, reflecting a relationship that shifted across decades rather than a stable structural connection. These findings imply that competitive balance concerns cannot be reduced to payroll disparities alone, as differences in spending failed to generate consistent or predictable competitive advantages.

Across professional sports, institutional arrangements generate systematic departures from competitive wage setting. Fort and Quirk (1995) document how league structures, monopsony power, and restrictions on player mobility create durable gaps between marginal revenue product and observed earnings. Limits on free agency, draft mechanisms, and arbitration procedures restrict players outside options and reduce the elasticity of labor supply, producing wage outcomes that diverge from competitive benchmarks. These constraints are most binding for early career players, who face limited mobility and delayed access to market-based bargaining. As a result, compensation adjusts slowly to changes in productivity, and wage growth often lags increases in league revenues or team level financial performance (Kahn, 2000). Becker (1964) and Mincer (1974) provide the human capital foundation for these patterns by outlining how experience, tenure, and skill accumulation shape wage trajectories over the career cycle, creating predictable earnings profiles that interact with institutional constraints in professional sports. Broader labor market effects associated with professional sports are also far weaker than commonly assumed. The presence of major league franchises and publicly financed stadiums does not raise local wages, and earnings in several service sector occupations are often lower in cities with professional teams. Expected spillovers from sports-related economic activity fail to materialize, as increases in consumer spending are offset by shifts from other local sectors rather than net expansion. The resulting wage patterns reflect displacement and reallocation rather than broad based economic gains, indicating that professional sports generate limited benefits for the wider labor market (Coates and Humphreys, 2001). Evidence from Ehrenberg and Smith (2017) points to wage suppression arising from monopsony power and institutional constraints, even in settings with high productivity and transparent output. Limited mobility, search frictions, and employer concentration weaken workers' bargaining positions and allow wages to fall below marginal revenue product. These mechanisms operate even when performance is easily observed and productivity differences are well documented, indicating that informational clarity alone does not guarantee competitive compensation. The patterns described in their work align closely with labor market conditions in professional sports, where restricted mobility and league level governance structures generate wage outcomes that diverge from competitive benchmarks despite precise measurement of individual output.

Salary determination in baseball reflects both player performance and the incentive structures created by league level financial rules. Revenue sharing arrangements alter the marginal return to winning, reducing the incentive for high revenue clubs to invest aggressively in talent and limiting the extent to which financial advantages translate into higher payrolls (Brown, 2009). Offensive statistics such as on base percentage and slugging percentage remain central predictors of pay, but their valuation is broadly uniform across franchises, indicating that differences in compensation arise from institutional incentives rather than divergent performance metrics. Salary equations are also sensitive to positional heterogeneity, as distinctions among starting, relief, and closing pitchers produce divergent career trajectories and compensation patterns, and similar differences extend across other positions where defensive responsibilities, offensive expectations, and career longevity vary substantially (Krautmann, Gustafson, and Hadley, 2003). Modeling player compensation therefore requires explicit attention to league level financial incentives and positional structures, since both dimensions condition how performance is translated into pay.

The expansion of data driven evaluation reshaped the understanding of salary determination in baseball. The Moneyball hypothesis, popularized by Michael Lewis and later tested econometrically by Hakes and Sauer (2006), argued that the market systematically undervalued certain skills, most notably on base percentage, while overvaluing traditional statistics such as batting average and runs batted in. Billy Beane and the Oakland Athletics applied this approach from 1999 through 2003, targeting players whose contributions were hidden in overlooked metrics and assembling competitive rosters despite limited payrolls. Hakes and Sauer confirmed that these inefficiencies were real in the late 1990s and early 2000s, and teams able to recognize undervalued skills gained a measurable advantage until the market corrected itself as other franchises adopted similar analytic approaches. Improved measurement not only altered how talent was priced but also exposed the fragility of salary structures built on conventional statistics. Fixed revenue streams, particularly league wide television contracts, further compress the distribution of salaries by weakening the link between individual performance and marginal revenue product, producing flatter compensation profiles than competitive models predict (Berri, Leeds, and von Allmen 2015). Modeling player compensation therefore requires attention to both evolving valuation methods and the revenue structures that condition salary outcomes, as these dimensions ultimately shape the translation of performance into pay.

Performance remains the foundation of pay determination, while the salary structure in MLB reflects the combined influence of athletic output, institutional mechanisms, and team financial capacity. The model incorporates performance metrics such as WAR and WS alongside measures of experience and team valuation to capture both the economic and structural dimensions of player compensation.

Salary determination in professional sports reflects the interaction between player productivity and the institutional structures governing labor markets. Evidence from the National Basketball Association indicates many players receive compensation below their marginal revenue product. Berri, Brook, and Schmidt (2007) estimate exploitation ratios and demonstrate that wage suppression is systematically produced by league rules that restrict mobility and delay access to competitive bargaining. In Major League Baseball, wage dispersion within teams carries measurable performance consequences. Depken (2000) documents a negative association between intra team salary inequality and on field outcomes, with performance declining as compensation becomes more uneven. Restrictions on player movement, including arbitration systems and delayed free agency, further limit bargaining leverage during early career and peak productivity years. Revenue concentration among a small number of teams reinforces employer influence and shapes the distribution of pay. Compensation outcomes therefore emerge from institutional constraints rather than pure productivity signals, and pay for performance mechanisms operate within these conditions, producing salary patterns anchored in bargaining structures instead of individual output alone.

Prior research has shown sabermetric measures such as WAR and related performance indicators correlate with salary, yet important gaps remain in understanding how teams translate on-field production into financial value. This study addresses these gaps by assembling a custom WS dataset covering multiple seasons, creating a consistent basis for comparing the latter with WAR and evaluating how each relates to compensation outcomes. The analysis introduces a structural modeling approach that views MLB's institutional rules as constraints within a financial optimization setting, reflecting how arbitration eligibility, free agency thresholds, and revenue-sharing mechanisms shape a valuation environment distinct from competitive labor markets. The discussion moves from the traditional labour-economics focus around the extent to which pay aligns with performance toward a financial examination of the implicit valuation model teams apply and the ways contract terms such as length, signing bonuses, and organizational financial capacity shape that model. This perspective highlights how risk, liquidity, and bargaining power affect pricing in ways econometric correlations alone cannot capture, offering a more complete understanding of how player salaries emerge within MLB's institutional and financial structure.

Research in sports labor markets has progressed from broad comparative analyses to increasingly granular, data intensive study of salary formation. Earlier contributions established the relevance of productivity measures, institutional constraints, and financial disparities across leagues, creating a foundation for more precise empirical work. As richer datasets and advanced performance metrics became available, researchers began to model compensation with greater detail, linking salary outcomes to measurable output, contract structure, and team level financial conditions. The resulting development points toward contemporary approaches that integrate modern analytics with econometric design to capture the full complexity of player valuation.

Recent contributions reflect a movement toward data rich, integrative modeling of salary determination. Simmons (2022) reviews developments in professional sports labor markets and highlights how research on pay performance relationships, discrimination, and mobility constraints has evolved alongside improvements in data availability and empirical methods. The progression toward more granular measurement and stronger identification strategies sets the stage for contemporary work that incorporates detailed performance metrics, contract structures, and institutional features into unified econometric frameworks. Butler, Farnell, and Simmons (2024) apply hierarchical models, fixed effects specifications, and nonlinear wage equations to detailed on pitch performance measures, contract terms, institutional rules, and club level financial variables in the English Premier League and Serie A. Salary levels in that setting vary with both individual output and structural conditions, with payroll constraints and league regulations altering the marginal effects of performance, and the multiple level structure separating within club and between club components of wage variation. Although the sport and institutional environment differ from MLB, the study provides a template for combining rich performance data, contract features, and financial constraints within a unified econometric framework for salary determination.

The evolution of research in this field reflects progression from theoretical models of performance-based pay and superstar effects to institutional analyses of labor market structures and, more recently, to data driven approaches integrating detailed player metrics with financial and contractual information. The cumulative evidence points to salary outcomes shaped by measurable on field productivity, experience based bargaining power, and the economic resources of individual teams. The dataset compiled for the analysis merges performance indicators, contract information, and team valuation into a coherent framework for assessing compensation variation across the league.

Data

An assessment of how MLB players are compensated relative to their performance requires a dataset that is both comprehensive and internally consistent. The analysis relies on a custom dataset that integrates detailed performance metrics with financial information drawn from established sources. Player statistics are collected primarily from Baseball-Reference, which providing season-level records of on-field performance and partial salary coverage. Measures of Wins Above Replacement (WAR) are included from both Baseball Reference and FanGraphs, yielding bWAR and fWAR, respectively. Contract structures and annual compensation figures are obtained from Spotrac, a repository of financial data in professional sports. The dataset also incorporates WS, following the framework introduced by Bill James and Jim Henzler in 2002. Together, these sources produce a unified, season-by-season player-level dataset that forms the empirical foundation of the study.

The dataset covers MLB seasons from 2015 through 2023. Beginning the sample in 2015 permits observation of player valuation immediately preceding the 2016 collective bargaining agreement, while the endpoint of 2023 reflects the most recent season with complete and internally consistent data across Baseball-Reference, FanGraphs, and Spotrac. Seasons beyond 2023 were excluded as several performance and salary fields remained incomplete or subject to revision at the time of data collection, preventing the construction of a fully reliable dataset. The resulting nine-year window spans two collective bargaining cycles and captures contemporary patterns in player performance, contract structure, and compensation.

Baseball Reference serves as the primary source of performance data, providing statistics for all players in the sample. Player performance is documented through standard offensive, pitching, defensive, and baserunning measures, along with advanced metrics such as WAR, on-base plus slugging, slugging percentage, and fielding independent pitching. Although the platform includes partial salary information, its financial coverage is incomplete, which motivates the use of Spotrac for contract structures and annual compensation. This combination of sources supports a detailed evaluation of player contributions across offensive, defensive, and pitching dimensions.

WAR provides a unified measure of a player's total contribution by estimating the number of additional wins he generates relative to a readily available replacement level alternative. The metric translates the diverse components of player performance, including hitting, baserunning, pitching, and fielding, towards a common currency of runs, reflecting the principle that individual actions matter only to the extent that they influence a team's ability to outscore its opponent. In practice, WAR aggregates three core elements: runs created by batters, runs prevented through pitching, and runs prevented through fielding. These components are combined within a single framework that allows direct comparison of players across positions and roles and offers a standardized assessment of marginal on field value.

Although WAR provides a unified measure of player value, bWAR and fWAR differ in how they convert performance into runs and wins. Hitting inputs remain essentially the same across the two systems; variation in WAR arises primarily from differences in pitching and fielding valuation. For pitchers, Baseball Reference uses a results based framework that incorporates the actual run outcomes of balls in play and the sequencing of events contributing to runs allowed, while FanGraphs employs a skills based approach that isolates outcomes under the pitcher's direct control, such as strikeouts, walks, hit batters, and home runs, and treats the remaining variation as a function of defense and the run environment. This distinction produces systematic differences in pitcher valuation. Pitchers who suppress runs despite modest strikeout rates often receive higher values in bWAR, while pitchers with strong strikeout profiles and limited contact quality tend to receive higher values in fWAR. Fielding also differs between the systems: Baseball Reference uses defensive runs saved to quantify defensive performance by comparing observed play outcomes to expected results, while FanGraphs relies on ultimate zone rating to estimate the likelihood a fielder should convert balls hit into specific zones into outs and evaluate performance relative to that benchmark. These defensive models do not always agree on the difficulty or run impact of specific plays, therefore fielders with strong range or atypical positioning can appear more valuable in one system than the other, with the differences most pronounced for players whose value is driven heavily by defense. Differences in pitching and defensive modeling generate systematic divergence in overall WAR estimates across the two systems. Interpreting these values follows a consistent scale: higher WAR reflects greater on field contribution, with 0 WAR denoting replacement level performance, approximately 2 WAR indicating a solid regular, 4 to 5 WAR corresponding to All Star caliber production, and 6 or more WAR representing MVP level value.

Win Shares are included as an additional measure of player performance, following the framework introduced by James and Henzler (2002). A claim point represents the basic unit used to allocate portions of a team's wins to individual players before those allocations are converted into WS. The metric distributes credit across offensive, defensive, and pitching contributions, and two variants of this measure are used in the dataset. First, Win Shares (WS) retains the original claim-point values, including negative values, allowing negative contributions to be recorded when a player reduces the team's expected contribution to winning. Second, Win Shares Claim Points Adjusted (WS CP Adj) applies a non-negativity constraint in which all claim-point values below zero are set to zero, reflecting the interpretation that a player cannot contribute negatively to a win. Including both variants allows the analysis to evaluate how alternative treatments of negative performance affect empirical estimates of player valuation.

Salary and contract information are collected from Spotrac for all player-season observations from 2015 through 2023. The dataset includes detailed information on salaries, bonuses, contract length, and free agency status, as well as reports on luxury tax obligations, a distinctive component of MLB's financial structure. Compensation in MLB is highly skewed, with a small subset of elite players earning disproportionately large salaries, reflecting the well-documented superstar effect that characterizes many performance-driven labor markets. This distributional pattern is shaped in part by MLB's institutional environment: unlike leagues such as the National Hockey League and the National Football League that impose hard salary caps, MLB operates under a luxury tax system that penalizes teams whose payroll exceeds a designated threshold. High-revenue teams frequently surpass this threshold and absorb the associated penalties, while lower-revenue teams may be constrained by it. Although these institutional features influence team-level spending behavior, the analysis in this study centers on player-level valuation rather than organizational financial strategy.

A central institutional feature shaping the salary data in this study is MLB's service-time–based compensation structure, which operates through salary arbitration and free agency. Salary arbitration eligibility typically arises after three to six years of service, during which the player and club submit proposed salary figures and an independent arbitrator selects one based on comparative performance and precedent. This process produces discrete adjustments in compensation that reflect both performance and institutional rules rather than market competition alone. Free agency, which generally becomes available after six full seasons, allows players to negotiate with any team, introducing market-driven salary variation and the potential for substantial increases in compensation. These mechanisms generate systematic differences in observed salaries across players with similar performance levels but different service-time classifications, making them an essential component of the compensation environment represented in the dataset.

Major League Baseball positions differ substantially in functional responsibilities, physical demands, and patterns of in-game usage, all of which contribute to systematic variation in observed performance and compensation. Pitchers are divided into starters and relievers, with starters carrying heavier workloads and relievers deployed in short, high-leverage situations. Catchers perform defensively intensive tasks and typically experience reduced playing time due to the physical demands of the position. Infielders vary in defensive range and throwing requirements, with shortstops and second basemen occupying high-mobility roles, third basemen emphasizing reaction time and arm strength, and first basemen generally facing lower defensive intensity but higher offensive expectations. Outfielders differ in spatial coverage, with center fielders responsible for the largest defensive area and corner outfielders more frequently evaluated on arm strength and offensive production. The designated hitter (DH), used universally beginning in 2022, serves as an offense-focused role in which a player bats in place of the pitcher and does not participate in defensive play. The DH was originally introduced in the American League in 1973 to address declining offensive output and to enhance spectator interest by replacing the pitcher's typically weak batting performance with a more productive hitter. Although often associated with offensive specialists, the role is frequently filled by players who also field other positions, as managers rotate athletes into the DH slot to provide partial rest or to accommodate short-term roster needs. As a result, DH assignments reflect a combination of offensive specialization, workload management, and roster flexibility. Managerial strategies, including rest-day rotations and matchup-based substitutions, further influence playing time across positions. Injury patterns also vary by role, with pitchers experiencing higher rates of arm-related absences and catchers facing elevated physical wear, leading to fluctuations in seasonal workload and performance opportunities. These positional and usage dynamics shape the distribution of performance statistics in the dataset and inform the interpretation of salary outcomes.

A distinctive feature of this dataset is the inclusion of WS, a performance valuation metric developed by Bill James unavailable in standardized databases as it requires a multistep calculation that allocates team wins across hitting, fielding, and pitching. To incorporate WS into this analysis, the metric is manually computed following James's original framework, ensuring a detailed and consistent assessment of player value.

WS quantifies a player's overall contribution to team success by converting individual performance into shares of team wins. Each win generates three shares, meaning a 90-win team produces 270 to be distributed among its contributors. The total pool is first divided across hitting, fielding, and pitching, and then allocated to individual players according to their measurable contributions in each area. This structure ensures that the sum of individual allocations aligns exactly with team performance and provides a unified measure of value across all positions. For pitchers, the calculation incorporates innings pitched, which reflect workload and durability, as well as earned run average, a central indicator of run prevention. Lower earned run averages result in a larger share of pitching credit, signaling more effective performance in limiting opponent scoring. For position players, allocations are based on offensive production such as Runs Created, along with defensive performance and overall team success. Runs Created captures a player's direct contribution to scoring, while defensive evaluations account for plays that prevent opposing runs, allowing the metric to balance both sides of a player's on-field responsibilities. By integrating these components, the metric offers a comprehensive assessment of a player's overall impact rather than emphasizing isolated statistics.

Player performance and salary data are matched across all sources using consistent identifiers derived from player names, team affiliations, and season records. Observations involving mid-season team changes or injury-related absences are treated with specific adjustments to ensure that both performance and compensation are accurately represented.

The final dataset contains 13,905 player-season observations from the 2015 through 2023 MLB seasons, including 21,600 team-level statistics across 80 variables for 30 teams. These data support the computation of park factors and the evaluation of team performance. At the player level, the dataset comprises roughly 115,000 batting statistics across 31 variables for 11,500 players, 120,000 pitching statistics across 34 variables for 9,000 players, and 155,000 fielding statistics across 27 variables for 15,500 players.

The analysis incorporates several player-level variables that document key aspects of career progression, contract structure, and demographic characteristics. These measures complement the performance metrics by capturing dimensions of player valuation that influence compensation independently of on-field output. Experience records accumulated MLB service time and reflects a player's position within the league's salary-determination framework, including arbitration eligibility and free-agency status. Years to free agency, derived from remaining contract years, provides information about a player's proximity to renegotiation opportunities and the strategic incentives associated with contract cycles. Age is considered to capture a player's stage within the typical career arc, while signing bonus records guaranteed upfront compensation at the time of entry into the league and reflects early-career expectations and prospect valuation.

Age is considered alongside experience to capture career-cycle effects that are not reflected in accumulated service time. Although the two variables are correlated, they represent distinct components of player valuation. Experience records progression through MLB's service-time system and the associated institutional salary adjustments, while age reflects expected durability, physical peak, and the timing of performance trajectories. Players with identical experience may differ substantially in age due to variation in debut timing, and players of similar age may have accumulated different levels of service time. Considering both variables allows the model to account for institutional and biological dimensions of compensation without attributing their combined effects to a single measure.

Signing bonus is considered to capture early-career valuation and the financial structure of initial contracts. The variable records guaranteed upfront compensation at the time of signing and provide information about prospect status, organizational expectations, and the allocation of risk between players and teams. Because signing bonuses are negotiated independently of annual salary and often preceded arbitration eligibility, they document an additional dimension of compensation that is not reflected in performance, experience, or age. Considering this measure allows the analysis to account for variation in initial contract terms that may influence subsequent salary outcomes.

At the team level, franchise valuation serves as a financial indicator sourced from Forbes. This measure integrates revenue, operating income, market size, brand strength, and historical transaction data, offering insight into the broader economic environment in which salary negotiations occur. Higher-valued franchises typically possess greater financial flexibility, potentially influencing both contract structures and player-level compensation outcomes.

The integrated dataset is structured for both cross sectional and panel data analysis, enabling the examination of salary–performance relationships while accounting for career dynamics, team strategies, and broader economic trends in MLB. By incorporating both WAR and WS, the dataset offers complementary perspectives on player valuation and supports robustness checks across alternative performance frameworks.

Methodology

An econometric framework is employed to analyze the relationship between player performance and salary compensation in MLB. The analysis focuses on the extent to which variation in on-field productivity corresponds to differences in individual earnings once experience, positional responsibilities, and institutional features of the league are incorporated. Regression models are estimated to quantify the marginal returns to performance, allowing productivity measures, career progression, and positional specialization to jointly inform salary outcomes.

Multiple linear regression models are estimated to evaluate the relationship between on field productivity and salary outcomes while incorporating experience and positional responsibilities. Because compensation data exhibit substantial dispersion, salaries are modeled in both levels and natural logarithms, the latter providing a more stable distribution and enabling coefficient estimates to be interpreted as approximate percentage changes in pay. In the level specifications, coefficients represent absolute dollar changes in salary associated with a one-unit change in performance. Robust standard errors are employed to account for potential heteroskedasticity in the error structure, ensuring that statistical inference remains reliable across the range of observed compensation levels.

Performance metrics form the core of the empirical framework, serving as the primary explanatory variables across all model specifications. Two measures are utilized: Wins Above Replacement (WAR) and Win Shares (WS). Each metric is incorporated in separate model specifications, allowing the analysis to evaluate compensation responses under alternative valuation approaches and to assess the robustness of results across distinct performance constructs. The relationship between performance and compensation is estimated using both linear and log-linear specifications, enabling comparisons across functional forms and allowing the analysis to account for differences in scale and distributional properties of salary data.

The baseline specifications are defined as:

$$Salary_i = \beta_0 + \beta_1 WAR_i + \varepsilon_i$$

$$Salary_i = \beta_0 + \beta_1 WS_i + \varepsilon_i$$

Corresponding log-linear models are:

$$\ln(Salary_i) = \beta_0 + \beta_1 WAR_i + \varepsilon_i$$

$$\ln(Salary_i) = \beta_0 + \beta_1 WS_i + \varepsilon_i$$

where,

$Salary_i$ represents total reported compensation for player i;

WAR_i denotes Wins Above Replacement;

WS_i denotes Win Shares, and

ε_i captures unobserved factors affecting compensation.

Extensions of the baseline models incorporate additional explanatory variables that capture established determinants of salary variation:

Experience is measured as the total number of completed MLB seasons prior to the observation year, reflecting accumulated human capital, including skill development, game knowledge, and consistency. Because experience also interacts with institutional features such as arbitration eligibility and free agency, its inclusion helps isolate the independent contribution of performance to compensation:

$$Salary_i = \beta_0 + \beta_1 Performance_i + \beta_2 Exp_i + \varepsilon_i$$

To allow the marginal return to performance to vary over a player's career, an interaction term between performance and experience is introduced. This specification accommodates the possibility that identical performance levels may be valued differently for early-career players compared with veterans:

$$Salary_i = \beta_0 + \beta_1 Performance_i + \beta_2 Exp_i + \beta_3 (Performance_i \times Exp_i) + \varepsilon_i$$

Positional indicators are included to account for systematic salary differences associated with role-specific skill requirements, defensive responsibilities, and league rules. Pitchers serve as the reference category, with all other positions represented through dummy variables:

$$Salary_i = \beta_0 + \beta_1 Performance_i + \sum_j \gamma_i Position_{ij} + \varepsilon_i$$

where,

$Position_{ij}$ is an indicator equal to 1 if player i appears at position j during the season, 0 otherwise.

Additional controls capture contractual and organizational factors. Contract length, measured as the number of years remaining on a player's current agreement, reflects the influence of guaranteed compensation and multi-year deal structures. Team value, expressed in billions of dollars, proxies for the financial capacity and market size of the player's club, allowing the model to account for systematic variation in salary offers across franchises.

All models are estimated using ordinary least squares. Given the substantial dispersion in salary data and the possibility that residual variance increases with compensation levels, heteroskedasticity-robust standard errors are employed to ensure valid inference across specifications. For models that include positional indicators or interaction terms, the same estimation framework is applied to maintain comparability across functional forms. The use of both level and log-linear specifications provides a robustness check on functional form assumptions, allowing the analysis to assess whether the estimated marginal effects of performance are sensitive to scale or distributional properties of the dependent variable. These components support a systematic evaluation of the determinants of salary variation.

To verify the accuracy of the estimation procedures, the full regression workflow was tested using the benchmark dataset provided by Petersen (2009). This dataset includes published coefficient and standard error values that serve as a reference for evaluating econometric routines. Replicating the benchmark regressions in R produced identical estimates and variance–covariance outputs, confirming that the data processing steps, model specifications, and heteroskedasticity-robust standard error calculations were implemented correctly. This validation step ensures that the MLB salary models are estimated within a verified and reliable computational framework.

The log transformation of salaries helps address the pronounced right-skewness of earnings in Major League Baseball and allows coefficients to be interpreted in percentage terms. Using both level and log-linear specifications provides estimates that are comparable across functional forms. Interaction terms between performance metrics and years of experience are included to allow the marginal effect of productivity to vary with seniority, consistent with the structure of MLB's labor market. Positional fixed effects are incorporated to account for systematic differences across roles, with pitchers serving as the reference category. Heteroskedasticity-robust standard errors are applied to support consistent inference throughout the analysis.

WS are incorporated into the analysis following the methodology developed by Bill James, providing a comprehensive measure of a player's total contribution to team success. The metric allocates each team's wins across offense, pitching, and defense, and then distributes these components to individual players based on their measurable on-field contributions. Offensive value is derived using the Runs Created framework, which converts a player's offensive performance into an estimate of total run production. Pitching value is based on run-prevention performance scaled by workload, while defensive value incorporates fielding outcomes adjusted for positional difficulty. Each team is assigned three WS per win, ensuring consistency and comparability across players, seasons, and clubs. In the empirical analysis, WS enter the regression specifications as an alternative performance metric to WAR, allowing the study to assess whether compensation responds similarly across distinct valuation frameworks and to evaluate the robustness of estimated performance–salary relationships.

The models incorporate a measure of player experience, defined as the number of completed Major League Baseball seasons prior to the observation year. Experience functions as a proxy for accumulated human capital, capturing variation in skill refinement, consistency, and game knowledge associated with tenure. It incorporates institutional features of MLB's labor structure, including salary arbitration and free agency, which generate systematic shifts in compensation opportunities across the career cycle. Arbitration eligibility typically arises after several seasons of service, while free agency generally occurs after six seasons, creating discrete changes in bargaining power and salary determination that are not exclusively tied to on-field performance.

These institutional mechanisms introduce additional variation in earnings that must be accounted for when estimating the relationship between performance and compensation.

To allow for heterogeneity in the returns to performance across career stages, selected model specifications include an interaction term between experience and WAR or WS. This structure allows the marginal effect of performance on salary to vary with tenure, acknowledging that equivalent performance levels may be compensated differently at early-career, mid-career, and veteran stages. Experience enters the regressions in contemporaneous form, ensuring that its influence is captured within each season's salary determination. Incorporating both experience and its interaction with performance yields a more comprehensive representation of how career progression, institutional rules, and productivity jointly shape compensation outcomes in MLB.

A separate dimension of salary variation from positional specialization is addressed by including categorical indicators for player positions. Positional roles in MLB vary in their defensive responsibilities, offensive expectations, and degrees of scarcity, generating structural variation in salary levels that are not captured solely by performance metrics. Pitchers serve as the reference category, and dummy variables are constructed for all other positions, including the designated hitter (DH), which became a league-wide role in 2022. Players who appear at multiple positions within a season are assigned to their primary position based on most defensive innings played, ensuring consistent classification across observations. Incorporating positional indicators mitigates omitted-variable bias arising from heterogeneous valuation across roles and allows the estimated returns to performance to reflect productivity rather than positional labor-market differences.

Observations are included in the analysis only when the underlying data are complete and free from reporting inconsistencies, ensuring reliability across the sample. Although the dataset is structured as a panel with repeated observations for individual players across seasons, the study employs a cross-sectional, season-level design in which each observation represents a player's performance and compensation each year. This structure facilitates a clear assessment of the contemporaneous relationship between productivity and salary while allowing the analysis to incorporate structural factors such as experience, contractual status, and positional responsibilities.

This methodological framework provides a comprehensive evaluation of how on-field performance, career experience, and positional role influence compensation outcomes. Cross-sectional, season-level analysis isolates contemporaneous effects, while the use of both linear and log-linear models captures absolute and proportional changes in pay. Interaction terms between performance metrics and experience allow the analysis to assess how the returns to productivity vary across career stages, reflecting both human-capital accumulation and institutional mechanisms such as salary arbitration and free agency. Taken together, these components establish a rigorous, multi-dimensional approach for examining the alignment between measurable performance and market compensation in professional baseball.

Results

The estimates across the models indicate stable associations between observable player attributes and compensation levels. The results show players with higher WAR tend to earn more once experience, positional role, and team characteristics are considered. These relationships appear consistently across datasets and functional forms, suggesting they capture persistent elements of teams' compensation practices. The coefficients can be read as reflecting the valuation weights teams assign to observable characteristics when structuring contracts, recognizing compensation emerges from the interaction of measurable performance with institutional and organizational conditions. Causality operates in both directions because teams seek high-performing players, shaping the salary distribution, and compensate them in ways influencing observed returns to performance.

The correlation matrix (Table 1) reports strong positive associations between the two WAR measures and between WAR and WS. WS and WS CP Adj are near-perfectly positively correlated. Experience yields weak correlations with all performance measures. Correlations are moderately positive between Contract Length and both WAR and WS, while team value has weak correlations with all variables. These patterns indicate that the performance metrics overlap but are not identical, justifying their examination in separate regression specifications.

Table 2 reports variance inflation factors for all baseline model specifications. Across every formulation, the VIF and adjusted GVIF values remain close to 1, indicating an absence of meaningful multicollinearity among the explanatory variables. Performance measures, whether based on WAR or WS, exhibit low VIF values, and the inclusion of Experience, Contract Length, and Team Value does not materially increase shared variance. The positional indicator, evaluated using the adjusted GVIF to account for its multiple degrees of freedom, also displays values near 1, confirming that positional controls do not introduce instability into the models. The measures suggest the regressors contribute distinct information and the estimated coefficients are not affected by linear dependence, ensuring the salary regressions are well conditioned and empirically reliable.

The baseline regressions (Table 3) evaluate the association between WAR and player salaries in Table 3. In the Spotrac models, a one unit increase in bWAR corresponds to approximately 1.451 million USD in additional salary, while the fWAR estimate is 1.628. The Baseball Reference models produce similar results, with coefficients of 1.491 for bWAR and 1.633 for fWAR. The estimates are tightly grouped across datasets and specifications, and the standard errors are small relative to the coefficients, reflecting precise and statistically strong relationships between WAR and compensation. In the Spotrac regressions, the adjusted R^2 values are 0.072 for bWAR and 0.087 for fWAR, with residual standard errors of 8.836 and 8.765. The Baseball Reference models show slightly higher explanatory power, with adjusted R^2 values of 0.092 for bWAR and 0.106 for fWAR, and residual standard errors of 7.949 and 7.889. These values show that WAR accounts for a limited share of salary variation, with substantial residual dispersion attributable to additional determinants such as experience, positional roles, contract structure, and team level financial conditions. The residual patterns reinforce this point. High performing players often earn salaries that exceed the predictions of a linear model based solely on WAR, while players with lower WAR values may still receive guaranteed minimums or incremental contractual increases. These outcomes reflect institutional features of MLB compensation, including arbitration rules, long term contract arrangements, positional scarcity, and variation in team valuation practices.

The second set of regressions in Table 4 incorporates experience alongside WAR. Across all specifications, experience is a large and precisely estimated determinant of salaries. In the Spotrac models, the coefficients are 1.277 for bWAR and 1.293 for fWAR, and in the Baseball Reference models they are 1.159 and 1.174. These values place the marginal effect of an additional year of service at roughly one million USD, conditional on performance. The similarity of the estimates across datasets establishes experience as a central component of salary determination. Two players with identical WAR receive different compensation when their MLB service time differs, consistent with the accumulation of tenure and the structure of the league's compensation system. WAR remains an important contributor to salary variation once experience is included. In the additive models, a one unit increase in WAR corresponds to 1.779 million USD in Spotrac and 1.789 million USD in Baseball Reference, with the alternative WAR measures at 1.922 and 1.963. When the interaction between WAR and experience is added, the direct effect of WAR declines to 0.706 in Spotrac and 0.763 in Baseball Reference, with the corresponding fWAR estimates at 0.792 and 0.963. The change in magnitude reflects the redistribution of variation once experience and its interaction with performance enter the specification. The interaction terms are positive and statistically significant in all models. The coefficients are 0.166 and 0.187 in Spotrac and 0.159 and 0.156 in Baseball Reference. These values establish that the salary return to performance increases with experience. A younger player and a veteran may produce the same WAR in a given season, but the veteran receives higher compensation for that output. This pattern aligns with the institutional features of MLB compensation, where early career salaries are constrained by arbitration rules and players with longer service time can secure higher pay for equivalent performance. Model fit increases once experience and the interaction terms are included. Adjusted R^2 rises from below 0.10 in the baseline WAR regressions to 0.321 and 0.342 in the Spotrac models and 0.341 and 0.360 in the Baseball Reference models when experience is added. With the interaction term, adjusted R^2 increases further to 0.334 and 0.357 in Spotrac and 0.354 and 0.372 in Baseball Reference. Residual standard errors decline across specifications, indicating that the expanded models account for a larger share of salary variation.

Table 5 present the subsequent regressions replace WAR with WS as the measure of player performance. The baseline models with only WS produce positive and precisely estimated coefficients. In the Spotrac dataset, the estimates range from 0.308 to 0.337, and in the Baseball Reference dataset they range from 0.310 to 0.336. When WS is used, the estimates fall between 0.328 and 0.350. These magnitudes are smaller than those obtained with WAR because WS is allocated in larger quantities across players, while WAR compresses value into fewer units. The difference in scale does not alter the underlying relationship, since both metrics assign higher salaries to higher levels of performance. The introduction of experience produces substantial changes in the estimates. Experience is strongly significant in all specifications, with coefficients between 1.0 and 1.2 million USD. These values place experience at a level comparable to the earlier regressions using WAR. The inclusion of experience also alters the intercepts. In the simple models with only WS, the intercepts are approximately six million USD, but they move toward zero or become negative once experience is added. This pattern reallocates variation initially appearing as a fixed salary level toward years of service, consistent with the institutional structure of MLB compensation. The interaction terms between WS and experience are positive and statistically significant in all models. The coefficients fall between 0.025 and 0.028. Although modest in size, these values accumulate over longer careers. Over a ten-year period, the marginal return to an additional WS increases by roughly 0.25 to 0.30 million USD. The estimates place the salary return to performance at a higher level for players with more experience, consistent with the compensation rules governing early career players and veterans. Model fit increases once experience is included. Adjusted R^2 rises from below 0.10 in the baseline specifications to values between 0.31 and 0.33 when experience is added. With the interaction term, adjusted R^2 increases further, reaching 0.319 in Spotrac and 0.335 in Baseball Reference. Residual standard errors decline across specifications, with the Baseball Reference models falling from 7.9 to 6.8. These changes place experience and its interaction with performance as central components of salary variation.

Table 6 reports regressions incorporating position and contract length, producing stable and precisely estimated coefficients on the performance variables. Across all specifications, bWAR ranges from 0.967 to 1.721 million USD in the Spotrac models and from 0.953 to 1.721 million USD in the Baseball Reference models. fWAR produces slightly larger estimates, with coefficients between 1.215 and 1.898 million USD in the Spotrac regressions and between 1.169 and 1.870 million USD in the Baseball Reference regressions. Values remain consistent with earlier models excluding position and contract length, and standard errors remain small across all specifications. Experience enters with large and precisely estimated coefficients. In the Spotrac models, estimates fall between 1.260 and 1.496 million USD, while in the Baseball Reference models they range from 1.140 to 1.381 million USD. Magnitudes remain similar across bWAR and fWAR specifications and remain stable once position and contract length are added. Standard errors, typically between 0.05 and 0.06 million USD, reinforce precision. Positional coefficients display substantial heterogeneity relative to pitchers. First basemen show positive estimates between 1.345 and 2.648 million USD in the Spotrac models and between 1.417 and 2.741 million USD in the Baseball Reference models. Third basemen also exhibit positive coefficients, ranging from 1.964 to 3.576 million USD in Spotrac and from 1.166 to 2.817 million USD in Baseball Reference. Outfielders and designated hitters show positive effects in several specifications, with DH coefficients reaching 3.109 million USD. Second basemen consistently display negative coefficients, ranging from 1.660 to 3.770 million USD across all models. Catchers and shortstops do not show systematic differences from pitchers, with coefficients small in magnitude and statistically insignificant in most specifications. Contract length enters positively and is precisely estimated in all models including it. Coefficients range from 1.284 to 1.382 million USD in the Spotrac regressions and from 1.319 to 1.382 million USD in the Baseball Reference regressions.

Standard errors, typically around 0.09 to 0.10 million USD, indicate high precision. Model fit improves once position and contract length are included. In the Spotrac regressions, adjusted R^2 increases from values between 0.335 and 0.359 in models without contract length to values between 0.420 and 0.435 in full specifications. In the Baseball Reference regressions, adjusted R^2 rises from 0.357 to 0.381 to values between 0.463 and 0.479. Residual standard errors decline across all models, with Spotrac specifications falling from approximately 7.48 to 7.34 million USD to values between 6.99 and 6.89 million USD, and Baseball Reference specifications falling from 6.69 to 6.56 million USD to values between 6.11 and 6.02 million USD.

Table 7 incorporates WS, WS CP Adj, experience, position, and contract length. Across all specifications, WS coefficients fall between 0.208 and 0.216 million USD in the Spotrac models and between 0.197 and 0.205 million USD in the Baseball Reference models. Standard errors remain small across all four regressions. Experience enters with large and precisely estimated coefficients, ranging from 1.482 to 1.491 million USD in the Spotrac models and from 1.367 to 1.376 million USD in the Baseball Reference models. Values remain consistent across WS and WS CP Adj specifications, and precision remains high throughout. Positional coefficients display variation relative to pitchers. First basemen produce small positive estimates in all models. Second basemen produce large negative coefficients, with values between 3.762 and 4.223 million USD in absolute magnitude across all specifications. Third basemen produce positive coefficients in the Spotrac regressions, with values near 1.8 to 1.9 million USD, while the Baseball Reference estimates remain smaller. Coefficients for catchers, outfielders, shortstops, and designated hitters remain close to zero with limited statistical significance across datasets and performance measures. Contract length enters positively and is precisely estimated in every specification. Coefficients fall between 1.405 and 1.451 million USD, with standard errors near 0.09 million USD. Adjusted R^2 values range from 0.422 to 0.465, and residual standard errors fall between 6.964 and 6.103 million USD. Values remain stable across datasets and across WS and WS CP Adj measures.

The positional patterns in the estimates indicate systematic differences in how teams compensate players across roles. Second basemen receive salaries below those of pitchers, third basemen receive higher salaries, and catchers experience the largest discounts. These relationships remain stable across performance metrics and suggest positional labor markets operate with distinct supply–demand conditions and structural valuation norms. Catchers face substantial physical demands, shorter expected career spans, and limited positional flexibility, which may lead teams to view compensation for this role as an investment with accelerated depreciation. Middle-infield positions combine high defensive responsibility with more modest offensive expectations, and teams may place greater weight on offensive production when determining compensation, producing lower salaries for second basemen. Third basemen receive higher salaries, consistent with the role's combination of moderate defensive requirements and elevated offensive expectations, as well as the relative scarcity of players who supply the level of offensive output teams seek at this position.

These positional effects persist after controlling for WAR and Win Shares, indicating teams incorporate structural role expectations and positional scarcity into compensation decisions independently of contemporaneous measured performance. The results align with an institutional view of the labor market in which salary reflects not only current output but also organizational assessments of positional risk, developmental investment, and opportunity cost.

The models summarized in Table 8 incorporate performance measures alongside team valuation, estimated both with and without experience. In the Spotrac specifications that exclude experience, performance variables enter with stable magnitudes. bWAR equals 1.366 million USD, and fWAR equals 1.551 million USD. WS and WS CP Adj produce coefficients of 1.416 and 1.564 million USD. Standard errors remain small across all four models. Team value contributes positively in each specification, with coefficients between 1.189 and 1.235 million USD. Adjusted R^2 values fall between 0.092 and 0.107, and residual standard errors range from 8.667 to 8.739 million USD. Adding experience alters the structure of the Spotrac regressions. Performance coefficients increase, with bWAR rising to 1.691 million USD and fWAR rising to 1.912 million USD. WS and WS CP Adj equal 0.323 and 0.335 million USD. Experience enters with coefficients between 1.216 and 1.298 million USD. Team value remains positive, with estimates between 1.240 and 1.318 million USD. Model fit improves once experience is included, with adjusted R^2 values between 0.335 and 0.366 and residual standard errors between 7.310 and 7.487 million USD. The Baseball Reference regressions excluding experience display similar patterns. bWAR equals 0.295 million USD, and fWAR equals 0.314 million USD. WS and WS CP Adj equal 0.298 and 0.317 million USD. Team value ranges from 1.057 to 1.109 million USD. Adjusted R^2 values fall between 0.112 and 0.126, and residual standard errors range from 7.801 to 7.864 million USD. Including experience increases the magnitudes of the performance variables in the Baseball Reference models. bWAR equals 1.712 million USD, and fWAR equals 1.891 million USD. WS and WS CP Adj equal 0.323 and 0.336 million USD. Experience enters with coefficients between 1.098 and 1.179 million USD. Team value remains positive, with coefficients between 1.103 and 1.184 million USD. Adjusted R^2 values rise to values between 0.348 and 0.384, and residual standard errors fall to values between 6.557 and 6.736 million USD. A further step in the analysis involves expressing salary in logarithmic form. Logarithmic specifications place salaries on a proportional scale and limit the influence of extreme contract values. The transformation follows standard empirical practice in earnings research and produces coefficients that reflect percentage changes in salary. The model structure remains unchanged from the linear specifications, with performance measures and team valuation entering first and experience added in a second group of regressions.

Table 9 reports the log specifications for the Spotrac data. In the models without experience, bWAR displays a coefficient of 0.174, corresponding to a 17.4% difference in salary for a one unit increase in bWAR. fWAR is measured at 20.6%, and WS and WS CP Adj take values of 4.1% and 4.4%. Intercepts range from 15.141 to 15.341, adjusted R^2 values fall between 0.068 and 0.100, and residual standard errors range from 1.077 to 1.096. These values remain stable across the four performance measures. Adding experience alters the structure of the Spotrac specifications by introducing a second set of log models that incorporate both performance measures and years of service. Experience ranges from 16.2% to 17.4%. Interaction terms remain near zero, with values of 0.001 for bWAR, –0.001 for fWAR, –0.001 for WS, and –0.001 for WS CP Adj. Adjusted R^2 values rise to the 0.357 to 0.366 range, and residual standard errors fall to values between 0.888 and 0.911. The Baseball Reference log models display similar patterns. In the regressions without experience, bWAR corresponds to a 17.8% difference in salary and fWAR to a 20.4% difference. WS and WS CP Adj take values of 4.1% and 4.4%. Intercepts range from 15.111 to 15.307, adjusted R^2 values fall between 0.072 and 0.102, and residual standard errors range from 1.064 to 1.081. Experience produces consistent changes in the Baseball Reference specifications. bWAR corresponds to a 22.0% difference and fWAR to a 25.2% difference. WS and WS CP Adj take values of 4.5% and 4.7%. Experience ranges from 15.9% to 16.6%. Interaction terms remain small, with values of 0.002 for bWAR, –0.0003 for fWAR, and –0.0005 for both WS variables. Adjusted R^2 values rise to the 0.345 to 0.372 range, and residual standard errors fall between 0.890 and 0.907.

Table 10 introduces the full log specifications that incorporate performance measures, experience, positional categories, and contract length. Across the eight models, intercepts range from 13.627 to 13.759, with standard errors near 0.073. Estimates are reported separately for Spotrac and Baseball Reference salary data. In the Spotrac bWAR specification, bWAR corresponds to a 15.0% difference in salary for a one unit increase in bWAR. Experience is measured at 18.8%, and contract length is set at 11.1%. Position indicators take values of –37.7% for second basemen and 24.2% for third basemen, while the remaining positions remain near zero. The adjusted R^2 equals 0.404 and the residual standard error equals 0.877. In the Spotrac fWAR model, fWAR is estimated at 19.3%. Experience remains near 18.7%, and contract length falls at 9.8%. Position indicators take values of 29.9% for third basemen, 17.0% for outfielders, and –25.9% for second basemen. The adjusted R^2 equals 0.431 and the residual standard error equals 0.856. Turning to the Baseball Reference salary data, the bWAR specification places bWAR at 15.0%, experience at 18.3%, and contract length at 11.0%. Position indicators take values of –35.0% for second basemen, 19.6% for outfielders, and 19.1% for third basemen. The adjusted R^2 equals 0.397 and the residual standard error equals 0.872. The Baseball Reference fWAR model reports fWAR at 18.7%, experience at 18.2%, and contract length at 9.9%. Position indicators take values of 27.2% for outfielders, 25.1% for third basemen, 28.0% for designated hitters, and –23.1% for second basemen. The adjusted R^2 equals 0.421 and the residual standard error equals 0.854. WS in the Spotrac specifications fall at 3.5% for WS and 3.7% for WS CP Adj. Experience remains near 18.7% and 18.5%, and contract length ranges from 11.7% to 11.4%. Position indicators take values of –46.8% and –46.1% for second basemen and values of 20.1% and 21.7% for third basemen. Adjusted R^2 values equal 0.416 and 0.419, and residual standard errors equal 0.868 and 0.866. In the Baseball Reference regressions, WS is measured at 3.4% and WS CP Adj at 3.6%. Experience spans 18.2% to 18.0%, and contract length is set at 11.7% and 11.5%. Position indicators remain similar to the Spotrac models, with second basemen at –43.5% and –42.8% and third basemen at 15.5% and 17.1%. Adjusted R^2 values equal 0.407 and 0.410, and residual standard errors equal 0.864 and 0.862. Across all eight specifications, performance measures, experience, positional categories, and contract length remain consistent in magnitude. Performance variables remain positive in every model, experience stays near 18%, and contract length ranges from 9.8% to 11.7%. Position indicators vary across roles but remain similar across datasets. Adjusted R^2 values fall between 0.397 and 0.431, and residual standard errors fall between 0.854 and 0.877.

Table 11 reports the log salary regressions that incorporate performance metrics, team valuation, and experience. The specifications extend the earlier models by adding franchise-level financial information, allowing performance and tenure to be evaluated alongside organizational resources. Results are presented for both Spotrac and Baseball Reference salary data and for each of the four performance measures, producing eight models per dataset. In the Spotrac bWAR models, bWAR is measured at 16.4% in the specification without experience, and team valuation takes a value of 13.9%. With experience included, bWAR rises to 20.8%, experience spans 17.1%, and team valuation is set at 14.6%. Intercepts range from 13.896 to 15.080, adjusted R^2 values span 0.086 to 0.377, and residual standard errors fall between 0.896 and 1.085. The Spotrac fWAR models follow a similar pattern. fWAR is estimated at 19.7% without experience and 24.5% with experience. Team valuation ranges from 13.8% to 14.5%, and experience takes a value of 17.3% in the full specification. Adjusted R^2 values range from 0.109 to 0.410, and residual standard errors range from 0.873 to 1.072. WS in the Spotrac dataset produce coefficients of 3.9% for WS and 4.2% for WS CP Adj in the models without experience, with team valuation ranging from 13.9% to 14.2%. With experience included, WS is measured at 4.3% and WS CP Adj at 4.5%. Experience spans 16.3% to 16.4%, and team valuation ranges from 15.1% to 15.3%. Adjusted R^2 values fall between 0.110 and 0.388, and residual standard errors fall between 0.889 and 1.071. The Baseball Reference bWAR models produce coefficients of 16.9% without experience and 21.0% with experience. Team valuation ranges from 12.5% to 13.2%, and experience takes a value of 16.4% in the full specification. Adjusted R^2 values range from 0.087 to 0.362, and residual standard errors range from 0.896 to 1.073. The Baseball Reference fWAR models produce fWAR coefficients of 19.5% without experience and 24.2% with experience. Team valuation ranges from 12.5% to 13.2%, and experience spans 16.6% to 16.7%. Adjusted R^2 values fall between 0.106 and 0.391, and residual standard errors fall between 0.878 and 1.062. WS in the Baseball Reference dataset produce coefficients of 4.0% for WS and 4.2% for WS CP Adj in the models without experience. With experience included, WS is measured at 4.3% and WS CP Adj at 4.5%. Team valuation ranges from 12.6% to 14.0%, and experience spans 15.6% to 15.8%. Adjusted R^2 values range from 0.110 to 0.371, and residual standard errors fall between 0.891 and 1.059. Across all specifications, the three variables of interest remain positive and statistically significant, with performance measures ranging from 3.9% to 24.5%, experience spanning 15.6% to 17.3%, and team valuation taking values between 12.5% and 15.3%. The explanatory power of the models varies across performance metrics and datasets, with adjusted R^2 values extending from 0.086 to 0.410 and residual standard errors falling between 0.873 and 1.085. The estimates form a consistent pattern across the sixteen regressions, with each component contributing a stable share of the variation in log salaries.

Table 12 reports salary regressions that incorporate interaction terms between performance metrics and team valuation. The eight specifications include four Spotrac models and four Baseball Reference models, each pairing a performance measure with team valuation and its interaction term. In the Spotrac WAR models, the bWAR specification produces coefficients of 10.7% for bWAR and 10.8% for team valuation, with the interaction term taking a value of 2.7%. The fWAR model yields estimates of 13.9% for fWAR, 10.0% for team valuation, and 2.9% for the interaction. Adjusted R^2 values for these models are 0.087 and 0.110, and residual standard errors are 1.084 and 1.071. The Spotrac WS models produce performance coefficients of 3.5% for WS and 3.8% for WS CP Adj. Team valuation ranges from 12.2% to 12.4%, and the interaction terms remain close to zero, with values of 0.2% in both specifications. Adjusted R^2 values are 0.110 and 0.117, and residual standard errors are 1.071 and 1.067. The Baseball Reference WAR models follow a similar structure. In the bWAR specification, bWAR is measured at 9.3%, team valuation at 8.5%, and the interaction term at 3.6%. The fWAR model produces coefficients of 11.5% for fWAR, 7.3% for team valuation, and 4.0% for the interaction. Adjusted R^2 values for these models are 0.090 and 0.109, and residual standard errors are 1.071 and 1.059. The Baseball Reference WS models produce coefficients of 3.2% for WS and 3.9% for WS CP Adj. Team valuation takes values between 9.8% and 11.5%, and the interaction terms remain small, ranging from 0.1% to 0.3%. Adjusted R^2 values are 0.110 and 0.116, and residual standard errors are 1.059 and 1.055. Within the eight specifications, the estimates remain positive and fall within narrow ranges. Performance coefficients span 3.2% to 13.9%, team valuation ranges from 7.3% to 12.4%, and the interaction terms take values between 0.1% and 4.0%. Adjusted R^2 values extend from 0.087 to 0.117, and residual standard errors range from 1.055 to 1.084. The results form a stable set of patterns across performance measures and salary sources, with the interaction terms contributing only limited variation relative to the main effects.

The previous regression models were estimated without age and signing bonus, as complete information on signing-bonus values is available only for a subset of players, resulting in a substantially smaller sample. To evaluate how these additional characteristics relate to salary outcomes, the extended analysis now considers both age and signing bonus within a consistent regression framework.

Table 13 summarizes the correlations among the variables incorporated into the extended regression framework. The addition of Age and Signing Bonus introduces two further relationships into the correlation structure. Age shows a moderate negative association with both bWAR (–0.201) and fWAR (–0.249), indicating that older players in the reduced sample tend to record lower recent performance metrics. Age is also strongly correlated with Experience (0.836), reflecting the mechanical link between career length and player age. Signing Bonus displays modest positive correlations with the performance measures, with values of 0.290 for bWAR and 0.251 for fWAR, and similar associations with both WS and WS CP Adj. Its correlation with Team Value (0.217) suggests that higher valued franchises are more likely to sign players who received larger bonuses. Neither Age nor Signing Bonus exhibits strong overlap with the remaining variables, allowing both to be incorporated into the extended regression specifications without concerns regarding multicollinearity.

Table 14 reports multicollinearity results using adjusted GVIF values. The adjusted GVIF values support the continued inclusion of Age and Signing Bonus, indicating that both variables contribute distinct information without introducing instability into the extended specifications. Incorporating these variables adds meaningful economic and contractual nuance to the salary determination framework. While performance, experience, positional role, and contract length remain central determinants, the addition of Age and Signing Bonus allows the analysis to more clearly separate career stage effects from the financial structure of compensation, strengthening the interpretive depth of the extended models.

The introduction of Age and Signing Bonus reduces the estimated effect of WAR while leaving the role of Experience largely unchanged, as reflected in Table 15. Across all specifications, Experience remains a strong and positive determinant of salary, with coefficients between 1.33 and 1.45 million USD per additional year. The inclusion of Age and Signing Bonus does not alter this pattern. The effect of WAR becomes smaller once these variables are added. In the bWAR models, performance is no longer statistically significant, and in the fWAR models the coefficient remains positive and significant at the 10 percent level but is smaller than in the baseline specifications. This indicates that Age and Signing Bonus absorb part of the variation previously attributed to on field performance. Relative to pitchers, the positional differentials remain stable after the inclusion of Age and Signing Bonus. First and third basemen continue to receive sizable premiums, generally between 3.7 and 4.6 million USD, while second basemen and catchers continue to show large and statistically significant discounts. Outfielders and shortstops retain positive but statistically weak coefficients, and designated hitters display estimates close to zero. Contract length continues to have a positive and highly significant effect, with annual premiums close to 0.91 to 1.00 million USD per additional year. The inclusion of Age and Signing Bonus does not change the magnitude or significance of this variable, indicating that contract duration remains a central component of salary determination. Age enters with positive coefficients between roughly 0.33 and 0.46 million USD per additional year, but it is only marginally significant in the Spotrac models and not significant in the Baseball Reference specifications. Its inclusion therefore does not materially change the structure of the models or the interpretation of the other variables. Signing Bonus is consistently positive and highly significant across all specifications. Each additional dollar of signing bonus corresponds to an increase in annual salary of roughly 1.38 to 1.65 dollars. This variable remains stable across all models and increases overall explanatory power, indicating that upfront financial commitments are strongly associated with higher annual compensation.

Table 16 provides the Win Share–based salary regressions, estimated with the same control variables used in the WAR specifications. Estimating the models with WS as the performance metric provides an alternative assessment of salary determination while holding experience, contract length, position, age, and signing bonus constant. This specification evaluates whether the earlier findings depend on the choice of performance measure. Experience continues to produce large and statistically significant coefficients in all Win Share models. The estimated return ranges from 1.35 to 1.46 million USD per additional year, closely aligned with the values obtained in the WAR-based specifications. The stability of these estimates indicates that the contribution of career accumulation to salary formation is not sensitive to the performance metric used. The Win Share variables themselves do not exhibit statistical significance. Neither WS nor WS CP Adj display meaningful coefficients in any specification, and the estimates remain small and imprecisely measured. This contrasts with the WAR-based models, where fWAR retained significance after the inclusion of Age and Signing Bonus. Once experience and contract structure are included, the Win Share measures add little explanatory power. Positional differentials relative to pitchers remain consistent with the earlier results. First and third basemen receive premiums between approximately 3.5 and 4.6 million USD. Second basemen show statistically significant discounts of roughly 4.0 to 4.5 million USD. Catchers display even larger negative coefficients, between 5.7 and 6.7 million USD. Outfielders and shortstops retain positive but statistically weak estimates, and designated hitters show no systematic salary advantage. The inclusion of Age and Signing Bonus does not materially alter these positional relationships. Contract length continues to yield strong and highly significant coefficients. Each additional contract year increases annual salary by approximately 1.01 to 1.02 million USD across all specifications. These magnitudes are nearly identical to those in the WAR-based models, indicating that the role of contract duration is stable across performance metrics. Age enters with positive but statistically weak coefficients. The estimated effect ranges from roughly 0.30 to 0.45 million USD per additional year, but the variable does not reach conventional significance levels in most models. This mirrors the WAR-based findings and indicates that age does not operate as an independent determinant of salary once experience is controlled for. Signing Bonus remains a consistently strong predictor of annual salary. Across all Win Share specifications, each additional dollar of signing bonus corresponds to an increase in annual salary between approximately 1.51 and 1.70 dollars, with high statistical significance. This confirms that signing bonuses are closely linked to long-term compensation rather than functioning as isolated upfront payments.

Table 17 presents the salary regressions that incorporate performance metrics together with team value, experience, and signing bonus. These models place the WAR- and WS-based measures within a common framework and apply the same set of controls to both Spotrac and Baseball Reference salary data, allowing the performance variables to be evaluated alongside indicators of organizational resources and contractual payments. Performance coefficients remain positive and statistically significant throughout the table. In the Spotrac regressions, bWAR and fWAR take values of 0.773 and 1.114, while the WS variables register coefficients of 0.165 and 0.167. The Baseball Reference estimates follow the same pattern, with bWAR and fWAR equal to 0.893 and 1.188, and the WS measures equal to 0.187 and 0.191. These magnitudes remain stable across salary sources and performance definitions. Team value contributes consistently across all eight specifications. The estimates range from 0.971 to 1.005 in the Spotrac models and from 1.018 to 1.051 in the Baseball Reference models. The narrow dispersion of these coefficients indicates a uniform association between franchise valuation and player salary across the different performance metrics. Experience shows a tightly clustered set of coefficients. Across all models, the estimates fall between 1.320 and 1.381 million USD per additional year, and each coefficient is highly significant. These values align closely with those in the earlier tables and remain unaffected by the choice of performance measure or salary source. Signing bonus enters each specification with a positive and strongly significant coefficient. The estimates range from 1.62 to 1.91, and the values remain nearly identical across the WAR- and WS-based models. The stability of these coefficients indicates that the inclusion of signing bonus does not alter the behavior of the other covariates. Adjusted R^2 values fall between 0.420 and 0.457 across the eight models. These values are comparable to those in the earlier tables and reflect consistent explanatory power once performance, team value, experience, and signing bonus are included.

The log-salary models in Table 18 incorporate performance metrics together with experience, position, contract length, age, and signing bonus for both Spotrac and Baseball Reference salary data. The estimates form a consistent structure across the two salary sources and four performance measures. Performance enters positively in all WAR specifications. In the Spotrac models, bWAR and fWAR take values of 0.043 and 0.072, while the Baseball Reference regressions produce coefficients of 0.049 and 0.071. These values show that the WAR measures contribute measurable positive components to log salaries across datasets. None of the WS or WS CP Adj variables reach significance in either dataset, with estimates between 0.007 and 0.009, indicating that the WS metrics contribute only small and statistically weak components once the full set of controls is included. Experience remains highly significant throughout. The Spotrac coefficients fall between 0.132 and 0.135, and the Baseball Reference values range from 0.158 to 0.161. These magnitudes show that experience consistently occupies one of the largest portions of the log-salary variation across all eight specifications. Position coefficients follow a consistent pattern. Second basemen and catchers show negative and significant estimates in every specification, while third basemen display positive coefficients across all models, with significance in several cases. Outfielders produce small positive estimates, with significance in the Baseball Reference fWAR and WS CP Adj models. These patterns show that positional categories contribute stable directional components to salary levels even after performance and contract characteristics are included. Contract length produces positive coefficients in all eight regressions. The Spotrac estimates range from 0.039 to 0.051, and the Baseball Reference values fall between 0.029 and 0.040, with significance in several specifications. These values show that contract length contributes a consistent positive component to log salaries across performance measures and salary sources. Age contributes positively in the Spotrac models, with coefficients between 0.059 and 0.062. In the Baseball Reference regressions, the corresponding estimates fall between 0.030 and 0.033 and do not reach significance. This pattern shows that age contributes a measurable component in the Spotrac specifications but a smaller and statistically weak component in the Baseball Reference models. Signing bonus is positive and highly significant in every model. The Spotrac coefficients range from 0.131 to 0.147, and the Baseball Reference values range from 0.147 to 0.159. These magnitudes show that signing bonus consistently contributes one of the strongest non-performance components to log salaries across all specifications. Adjusted R^2 values fall between 0.502 and

0.534 across the eight specifications, showing that the combined set of performance, experience, position, contract length, age, and signing bonus explains a similar share of log-salary variation across datasets and performance measures.

The estimates in Table 19 summarize log-salary regressions that combine performance metrics with team valuation, experience, and signing bonus. These specifications narrow the focus to variables that operate at the player and franchise level, producing a compact set of coefficients across both salary sources. Across the Spotrac WAR models, performance contributes positive components to log salaries. The bWAR coefficient is 0.054, and the fWAR estimate rises to 0.088. Team value enters with coefficients of 0.128 and 0.133, and experience contributes values of 0.155 and 0.159. Signing bonus remains positive and highly significant in both cases. Together, these estimates show that performance, team valuation, experience, and signing bonus each supply stable contributions within the Spotrac WAR framework. Adjusted R^2 values range from 0.478 to 0.492. The Baseball Reference WAR models display a similar configuration. The bWAR and fWAR coefficients are 0.066 and 0.091, and team value contributes estimates of 0.124 and 0.130. Experience enters at 0.159 and 0.162, and signing bonus remains strongly significant. These values show that the four variables contribute consistently across the two WAR measures, with adjusted R^2 values of 0.504 and 0.513. The Spotrac WS specifications produce small but positive performance coefficients. WS and WS CP Adj each take values of 0.011. Team value contributes coefficients of 0.131 and 0.130, and experience enters at 0.155 in both models. Signing bonus remains positive and significant. These estimates show that the WS measures contribute modest components to log salaries, while team valuation, experience, and signing bonus maintain stable magnitudes. Both models yield an adjusted R^2 of 0.475. The Baseball Reference WS models follow the same pattern. WS and WS CP Adj each produce coefficients of 0.013, and team value contributes estimates of 0.128 and 0.127. Experience enters at 0.158 in both specifications, and signing bonus remains strongly significant. These values show that the WS measures and the three non-performance variables contribute consistently across the two Baseball Reference WS specifications, each with an adjusted R^2 of 0.499. Across all eight regressions, the coefficients for performance, team valuation, experience, and signing bonus remain tightly grouped, and the adjusted R^2 values fall between 0.475 and 0.513. The estimates form a uniform pattern in which each variable contributes a stable component to log salaries across performance metrics and salary sources.

Taken together, the estimates across all tables form a consistent empirical pattern. Performance, experience, team valuation, contract characteristics, age, and position each contribute measurable components to salary variation, and the magnitudes of these coefficients remain stable across datasets, performance metrics, and model forms. The linear and log-linear specifications produce closely aligned results, with adjusted R^2 values that fall within narrow ranges for each group of models. Across all regressions, the variables that enter significantly do so with similar signs and comparable magnitudes, and the non-significant variables remain limited to a small subset of performance and demographic measures. The full set of models therefore provides a coherent empirical account of the observable components associated with salary levels in Major League Baseball.

Discussion

The estimates across the models outline a compensation structure in which performance measures, service-time progression, and organizational characteristics each account for identifiable portions of salary variation. Productivity variables enter consistently, though their magnitude shifts with contract status and career stage. Team-level financial and structural differences also appear in the results, producing variation in how comparable levels of on-field output correspond to pay. The combined patterns depict a system where multiple observable factors operate concurrently, with no single component fully accounting for the distribution of salaries.

Across all specifications, WAR based measures consistently produced larger and more stable coefficients than the WS metrics. This pattern held in both level and log linear models and across both salary sources. The relative strength of WAR reflects its design as a marginal value measure anchored to a replacement baseline, which allows individual contributions to be expressed in units that remain comparable across positions and seasons. The additive structure of the measure and its direct correspondence to incremental wins align with the way teams evaluate performance when determining compensation. By limiting the influence of team context and reducing measurement error, WAR provides a more consistent signal of individual output in models where precision in the performance variable is essential. Stability across seasons and applicability across roles further reinforce the close relationship between this measure and the salary setting process.

The weaker and more variable coefficients on WS and WS CP reflect the structural properties of the framework itself. Retrospective allocation of team performance across all contributors disperses value and prevents marginal effects from being isolated. Its credit allocation rules incorporate assumptions about defensive context, team environment, and playing time distribution, making the measure more sensitive to contextual variation. This blending of individual and team influences increases the degree to which external factors shape measured value and can heighten correlations with other contextual variables in the model. As a result, WS and WS CP capture a broader conception of contribution but do so in a form that corresponds less directly to the mechanisms through which individual performance is translated into compensation.

The magnitude of performance coefficients declined as additional controls were introduced, particularly once experience, position, contract length, and team valuation entered the models. The observed changes in estimates are consistent with evidence where performance explains only part of salary variation and structural factors absorb a substantial share of what initially appears to be productivity driven. These controls reflect institutional features of the labor market, encompassing career earnings trajectories, role expectations, contractual risk structures, and financial headcount capacity in organizations. Because these dimensions influence compensation independently of on field output, their inclusion necessarily diminishes the share of salary variation attributed solely to performance. This pattern is consistent with the conclusions of Medoff and Abraham (1980), indicating compensation systems frequently reward factors unrelated to measurable productivity and that institutional structures guide earnings trajectories beyond what performance alone would predict.

The log salary models reinforced this interpretation, with performance retaining significance but accounting for a smaller share of explained variation once the full set of controls was included. The resulting pattern aligns with the dynamics outlined by Jensen and Murphy (1990), concluding compensation systems often incorporate risk considerations, organizational constraints, and incentive design features that limit the direct responsiveness of pay to output. As additional dimensions enter the specification, the remaining performance coefficient reflects a more precise estimate of its contribution within a compensation structure shaped by career progression, positional scarcity, contractual arrangements, and team level financial conditions. The results suggest productivity remains an important determinant of salary, although the influence of performance is moderated by the broader economic and institutional environment in which compensation decisions occur, a relationship consistent with prevailing views of how value is translated into pay in MLB.

Experience emerges as the most stable and economically meaningful variable across all models. Its coefficients remain large and positive across performance metrics, functional forms, and salary sources, aligning with human capital frameworks in which tenure reflects accumulated skill, knowledge, and reliability. Additionally, it corresponds to research on internal labor markets manifesting progress in compensation through institutional rules rather than current season performance. MLB's compensation structure reinforces this dynamic, as arbitration eligibility and free agency introduce discrete shifts in bargaining power tied to service time. These institutional features generate predictable salary progression independent of performance, rendering experience coefficients capturing both accumulation of productive attributes and structural mechanisms through which MLB allocates earnings across playing careers.

The interaction terms between performance and experience in the OLS specifications indicate the marginal return to performance increases with tenure, while the logarithmic specifications show minimal interaction effects. In the OLS framework, salary is modeled in dollar terms, allowing the estimates to capture absolute changes in compensation. The logarithmic framework models salary in proportional terms, allowing the estimates to capture percentage changes. These two formulations reveal different dimensions of how compensation responds to performance across the career. When salary is expressed in dollars, experienced players receive larger monetary gains from equivalent performance improvements, reflecting the higher salary baselines associated with arbitration eligibility, free agency status, and accumulated bargaining power. When salary is expressed proportionally, percentage changes remain comparatively stable across the career, indicating the presence of institutional constraints that limit the influence of performance once players enter established contractual tiers. Taken together, the results imply performance contributes more strongly to absolute earnings for experienced players, while proportional adjustments remain relatively uniform for both rookies and veterans, underscoring the extent to which MLB's compensation system shapes the interaction between performance and experience.

The relevance of institutional wage setting is consistent with literature, particularly Oyer and Schaefer (2011), studying compensation systems which incorporates tenure-based structures that influence earnings independently of output. MLB's arbitration and free agency rules create precisely such conditions, producing salary patterns shaped by service time, bargaining power, and contract structure. In this environment, experience coefficients reflect not only accumulated human capital but also the operation of an internal labor market in which compensation evolves through institutional rules, tenure based bargaining power, and long-term contractual arrangements. These mechanisms determine why observed salaries reflect a combination of productivity, risk allocation, and organizational constraints rather than performance alone.

Positional indicators introduce systematic and economically significant variation into salary outcomes, revealing a consistent hierarchy. Second basemen and catchers receive lower compensation compared to pitchers, while third basemen and, in several specifications, outfielders receive higher valuations. This conclusion persists across performance metrics and salary sources, encouraging a stable positional structure rather than model-specific fluctuations.

The consistency of these positional effects after controlling for performance, experience, and contract characteristics suggests positional labor markets operate with distinct valuation regimes. Earlier work in labor economics highlights persistent differences in how MLB organizations price defensive responsibilities, offensive expectations, and positional scarcity, and the current estimates reinforce this broader pattern. The results also align with Rosen's (1981) superstar framework, which emphasizes the tendency of markets to assign disproportionate value to roles with elevated strategic importance or limited supply even when output is observable. In suc setting, positional coefficients capture measurable productivity as well as the structural hierarchy embedded in MLB's allocation of value across defensive and offensive roles.

Contract structure exerts a substantial and systematic influence on salary determination as it enters with consistently positive coefficients across all specifications, indicating a clear premium associated with multi-year agreements even after controlling for performance, experience, and positional attributes. This dynamic reflects the intertemporal nature of long-term contracting in professional sports, where extended agreements shift risk toward teams, constrain future roster flexibility, and compensate players for expected rather than contemporaneous output. The persistence of this premium indicates teams incorporate forecasts of aging, durability, and role stability into present-day salary commitments. Signing bonus generates some of the largest and most stable estimates in the expanded models, underscoring the central role of guaranteed, upfront payments in the capitalization of total player value, suggesting signing bonuses as a strategic mechanism for converting expectations about future productivity into immediate, risk-free compensation insulated from performance volatility. The consistency of these effects indicates that guaranteed compensation functions as a primary mechanism through which teams' price both expected performance trajectories and the uncertainty embedded in long-horizon commitments.

Team valuation remains positive and significant across all models, indicating franchise-level financial capacity contributes independently to salary outcomes. This suggests broader bargaining sets, greater tolerance for long-term commitments, and a reduced sensitivity to performance volatility when negotiating contracts for financially stronger organizations. Such dynamic aligns with established evidence on revenue disparities and payroll stratification in MLB, where higher-valuation teams consistently sustain larger payrolls and exercise greater flexibility in contract design. A central implication is salary outcomes are jointly determined by individual productivity and franchise-level financial capacity, resulting in compensation structures systematically conditioned by the resource inequalities embedded within MLB's economic landscape.

The models in Tables 15–17 deliver a consistent pattern in the signing-bonus coefficient. Across the specifications that include this variable, higher signing-bonus amounts are associated with higher annual salaries, with the estimates supporting a multiplying factor between 1.5 and 1.7. This magnitude offers insight into how teams distribute compensation across different components of a contract. A signing bonus is a guaranteed, upfront payment that does not depend on future performance, injuries, or contract adjustments. Financially, it functions as a risk-free element of total compensation, whereas annual salary may vary with incentives, arbitration outcomes, or renegotiation. The multiplier observed in the estimates suggests teams attach a premium to guaranteed compensation, a pattern consistent with several mechanisms related to risk, commitment, liquidity, and bargaining dynamics.

Risk considerations arise because converting uncertain future performance into an upfront payment reduces exposure to performance volatility, availability, and organizational changes. Commitment effects may also be present, as larger bonuses can reflect organizational confidence in a player's projected contribution and provide stability regarding role and usage. Liquidity and timing considerations influence contract design as well, since teams with greater financial flexibility may prefer front-loaded structures, while others rely more heavily on incentive-based or back-loaded arrangements. Bargaining dynamics play a role because players with stronger negotiating positions, including free agents and arbitration-eligible veterans, tend to secure larger bonuses relative to salary.

The present analysis concludes, using player-level data, that team financial capacity exerts an independent and statistically robust influence on salary outcomes even after accounting for performance. This pattern infers the idea where compensation in MLB reflects not only individual productivity but the resource environment within which negotiations occur. The mechanism is consistent with the league-level financial disparities documented by Fort and Quirk (1995), who demonstrate how large-market and small-market franchises differ in their ability to sustain payroll investment. It aligns additionally with Kahn's (1993) emphasis on how institutional and financial conditions shape the bargaining set available to players across professional sports. Unlike these earlier structural and conceptual accounts, the present study provides direct econometric evidence at the player level, showing how the financial inequalities they identify manifest in individual compensation patterns.

The empirical results reveal a salary structure shaped by the interaction of productivity, institutional constraints, and organizational resources, a configuration anticipated across several strands of the literature. The relatively modest performance coefficients reflect the incentive compatibility limits accentuated by Lazear (1981) as pay is effective only when institutions allow wages to track output without excessive risk or administrative rigidity. Their attenuation is also consistent with the weak pay performance sensitivity documented by Medoff and Abraham (1980), where results define firms rely heavily on seniority rather than contemporaneous productivity; and by Jensen and Murphy (1990) studies revealing high-stake environments where marginal effect of performance on pay is marginally small once governance structures, risk aversion, and institutional constraints are considered. In MLB's institutional environment, where pre arbitration and arbitration rules restrict salary adjustments, performance fails to translate into compensation with the elasticity predicted by a pure marginal productivity model. The strong and stable experience effects in the present analysis reinforce this interpretation mirroring the tenure-based progression patterns identified in traditional labor markets while simultaneously reflecting the arbitration and free agency thresholds highlighted by Kahn (1993) and Oyer and Schaefer (2011) as arbitration and free agency create discrete shifts in bargaining power and compensation. These institutional discontinuities generate predictable salary jumps that operate independently of performance, absorbing variation that incentive theory alone is unable explain.

The distributional features of the data correspond closely to the mechanisms articulated in Rosen's (1981) analysis of superstar labor markets, though these dynamics emerge primarily within the upper segment of the MLB salary distribution rather than uniformly across players. Rosen demonstrates how scalable performance and broad audience reach create convex earnings profiles, where small differences in underlying talent produce disproportionately large differences in compensation. The pronounced right skew in MLB salaries and the substantial returns accruing to elite performers in the present analysis reflect this convexity with precision: top players operate in an environment where their performance is broadcast, monetized, and consumed at scale, enabling marginal improvements in productivity to generate outsized economic rewards. The empirical results reinforce this mechanism through sharply increasing performance coefficients in the upper quantiles, revealing a compensation structure where market forces exert considerable influence only once players reach performance levels high enough to activate Rosen's scalability conditions. The weak median pay performance sensitivity in the models indicates a contrasting regime in which most players remain subject to institutional constraints rather than market driven convexity. This dual structure produces a labor market where institutional wage compression governs early and mid-career earnings, while superstar dynamics shape compensation for a relatively small subset of elite performers, aligning the empirical evidence with Rosen's theoretical predictions in a highly differentiated manner

The persistent significance of team valuation introduces organizational capacity as an independent determinant of compensation. Salary advantages for players on financially stronger teams correspond to the institutional framework developed in Fort and Quirk's (1995) theoretical analysis, as the model links variation in local revenue potential and ownership resources to persistent differences in the financial strategies franchises employ when allocating labor expenditures. Even after controlling for performance, experience, and position, players on higher valuation teams receive systematically higher salaries, indicating financial resources shape bargaining outcomes in ways extending beyond productivity and institutional rules. The empirical results thus undervalue how structural disparities focused in Fort and Quirk's analysis manifest at the individual level, with valuation effects remaining positive, statistically significant, and economically meaningful across specifications. MLB's absence of a salary cap reinforces this dynamic, as franchises must distribute payroll across an entire roster; teams with substantial financial resources can accommodate the cost of elite contracts without compromising their ability to staff the remainder of the lineup, whereas teams with limited revenue capacity face binding budget constraints limiting access to high-cost talent. This environment produces a compensation structure influenced not only by incentive design and contractual institutions but also by the financial heterogeneity embedded in MLB's organizational architecture, where differences in revenue capacity translate into distinct bargaining environments across franchises.

The results point to several implications for how teams approach player valuation and capital allocation. The estimates indicate scope for a more formal risk-adjusted valuation framework, as long-term contracts are often shaped by institutional rules such as arbitration and free agency rather than by explicit discounted-cash-flow assessments of expected future performance. A more disciplined approach would incorporate projections adjusted for injury risk, aging profiles, and regression toward the mean, discounted at a rate reflecting a team's opportunity cost of capital. WAR and Win Shares could serve as inputs to such a framework, although their current use appears uneven across organizations, and greater consistency could improve allocation efficiency and reduce systematic mispricing. The signing-bonus multiplier observed in the estimates further suggests that teams place substantial value on guaranteed, upfront compensation. A signing bonus converts uncertain future performance into a certain present payment, functioning as a hedge against volatility in player availability and output. The magnitude of the multiplier implies teams may be paying a premium for this certainty relative to what a risk-adjusted valuation would imply, and a more refined approach would estimate the implicit volatility of player performance, assess the option value embedded in contract flexibility, and price guaranteed components accordingly.

The relationship between team valuation and individual salaries highlights the role of organizational finance in shaping labor-market outcomes. Wealthier teams appear able to attract higher-performing players and compensate them more generously, reinforcing competitive advantages over time. This dynamic has implications for league competitive balance, as financial disparities translate into differences in roster quality even after accounting for performance, experience, and position. In the absence of hard salary caps, these patterns suggest that financial strength naturally influences competitive outcomes. These insights have relevance for contract design, salary-cap planning, and prospect valuation, and they also bear on future collective bargaining discussions. Understanding how institutional rules shape earnings distributions may help clarify whether alternative mechanisms, including adjustments to luxury-tax thresholds or revenue-sharing arrangements, could better align compensation with underlying player value.

Conclusion

The analysis examined the determinants of MLB player compensation through an integrated econometric framework combining performance metrics, institutional progression, and organizational resources. The evidence indicates salary formation reflects a layered equilibrium in which productivity, experience, and team level financial capacity operate concurrently rather than as isolated determinants. The estimates provide player level confirmation of mechanisms typically examined separately in the labor economics and sports economics literatures, including incentive moderation under constrained wage responsiveness, the limited elasticity of pay with respect to performance, the convex returns accruing to elite contributors, the institutional structuring of earnings through tenure-based rules, and the influence of franchise level financial capacity. When operating within a unified labor market, the underlying dynamics combine to produce a compensation structure defined by their interaction rather than by any isolated mechanism. The resulting configuration reflects a system in which economic incentives, institutional constraints, and organizational heterogeneity jointly shape how value is translated into pay.

The empirical results highlight several central features of MLB's labor market. WAR aligns more closely with salary outcomes than Win Shares, which suggests marginal value metrics may play a more prominent role in contract negotiations than aggregate performance measures. Experience remains a dominant determinant of compensation, which reinforces the continued influence of institutional structures rewarding tenure independently of on field productivity. Positional differences persist even after controlling for performance, which indicates role specific valuation remains embedded in organizational decision making. Team financial capacity also contributes meaningfully to salary outcomes, which underscores the importance of organizational context in shaping the distribution of earnings.

Limitations arising from the data and empirical approach challenge the interpretation of the results by constraining the scope of the inferences that can be drawn. The cross-sectional structure of the dataset restricts the capacity to capture dynamic effects, player specific heterogeneity, or contract level adjustments over time. Performance metrics, although comprehensive, may not fully represent intangible contributions or role specific nuances. Reported salary data may exclude incentives, escalators, or deferred compensation, and franchise level valuation measures may not reflect short term liquidity constraints or internal budgeting practices. These limitations do not diminish the empirical patterns but indicate areas where future research could refine measurement and improve identification.

The analysis also suggests several directions for continued inquiry. Panel data would permit the estimation of player fixed effects and the modeling of dynamic contract adjustments. Nonlinear specifications could clarify whether marginal returns to performance vary across the talent distribution. Additional work could examine how incentive structures interact with performance metrics, or how organizational strategies differ across teams with varying financial capacities. Broader extensions could incorporate negotiation dynamics, risk preferences, or roster construction philosophies to deepen the understanding of how MLB organizations convert player value into compensation.

The results indicate MLB compensation reflects the combined influence of productivity, institutional rules, and organizational resources. The labor market does not align with any single theoretical framework. It exhibits a structure in which economic, institutional, and organizational forces operate jointly to determine how value is priced. The analysis advances that understanding by providing an integrated empirical account of salary formation and by identifying the mechanisms through which performance and context interact to shape earnings in professional baseball.

References

Berri, D. J., Brook, S. L., & Schmidt, M. B. (2007). *The wages of wins: Taking measure of the many myths in modern sport*. Stanford University Press.

Berri, D. J., Leeds, M., & von Allmen, P. (2015). Salary determination in the presence of fixed revenue streams: The case of Major League Baseball. *Journal of Sports Economics, 16*(1), 3–21.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.

Brown, R. W. (2009). Revenue sharing and competitive balance in professional sports leagues. *Journal of Sports Economics, 10*(1), 3–20.

Butler, R., Farnell, T., & Simmons, R. (2024). Hierarchical wage determination in European football: Performance, contracts, and financial constraints. *European Sport Management Quarterly*. (Forthcoming)

Coates, D., & Humphreys, B. R. (2001). The economic consequences of professional sports strikes and lockouts. *Southern Economic Journal, 67*(3), 737–747.

Depken, C. A. (2000). Wage disparity and team productivity: Evidence from Major League Baseball. *Economics Letters, 67*(1), 87–92.

Ehrenberg, R. G., & Smith, R. S. (2017). *Modern labor economics: Theory and public policy* (13th ed.). Routledge.

Fort, R., & Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature, 33*(3), 1265–1299.

Hakes, J. K., & Sauer, R. D. (2006). An economic evaluation of the Moneyball hypothesis. *Journal of Economic Perspectives, 20*(3), 173–185.

Hall, S., Szymanski, S., & Zimbalist, A. (2002). Testing causality between team performance and payroll: The case of Major League Baseball. *Journal of Sports Economics, 3*(2), 149–168.

Jensen, M. C., & Murphy, K. J. (1990). Performance pay and top-management incentives. *Journal of Political Economy, 98*(2), 225–264.

Kahn, L. M. (1993). Free agency, long-term contracts, and compensation in Major League Baseball. *Economic Inquiry, 31*(1), 1–16.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *Journal of Economic Perspectives, 14*(3), 75–94.

Krautmann, A. C., Gustafson, E., & Hadley, L. (2003). A note on the structural stability of salary equations: Major League Baseball pitchers. *Journal of Sports Economics, 4*(1), 56–63.

Lazear, E. P. (1981). Agency, earnings profiles, productivity, and hours restrictions. *American Economic Review, 71*(4), 606–620.

Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy, 89*(5), 841–864.

Medoff, J. L., & Abraham, K. G. (1980). Experience, performance, and earnings. *Quarterly Journal of Economics, 95*(4), 703–736.

Mincer, J. (1974). *Schooling, experience, and earnings*. Columbia University Press.

Oyer, P., & Schaefer, S. (2011). Personnel economics: Hiring and incentives. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 4B, pp. 1769–1823). Elsevier.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature, 37*(1), 7–63.

Rosen, S. (1981). The economics of superstars. *American Economic Review, 71*(5), 845–858.

Rosen, S., & Sanderson, A. (2001). Labour markets in professional sports. *Economic Journal, 111*(469), F47–F68.

Scully, G. W. (1974). Pay and performance in Major League Baseball. *American Economic Review, 64*(6), 915–930.

Simmons, R. (2022). Sports labor markets: Developments, debates, and directions. *Oxford Review of Economic Policy, 38*(4), 676–699.

Appendix A - Tables

Table 1. Correlation Matrix of Key Variables

Variable	bWAR	fWAR	WS	WS CP Adj	Experience	Contract Length	Team Value
bWAR	1.000						
fWAR	.834	1.000					
WS	.799	.811	1.000				
WS CP Adj	.799	.817	.992	1.000			
Experience	- .121	- .129	- .055	- .042	1.000		
Contract Length	.489	.440	- .320	.380	.382	1.000	
Team Value	.109	.097	- .025	.081	.087	.065	1.000

Note: WS CP Adj are Win Shares adjusted for negative claim points.

Table 2. Variance Inflation Factors for Baseline Models

Model Specification	Variable	VIF / GVIF	Df	Adjusted GVIF
Salary = bWAR + Experience	bWAR	1.015		
	Experience	1.015		
Salary = fWAR + Experience	fWAR	1.017		
	Experience	1.017		
Salary = bWAR + Experience + Position	bWAR	1.139	1	1.067
	Experience	1.065	1	1.032
	Position	1.175	7	1.012
Salary = fWAR + Experience + Position	fWAR	1.051	1	1.025
	Experience	1.064	1	1.031
	Position	1.082	7	1.006
Salary = bWAR + Experience + Position + Contract Length	bWAR	1.372	1	1.171
	Experience	1.168	1	1.081
	Position	1.206	7	1.013
	Contract Length	1.445	1	1.202
Salary = fWAR + Experience + Position + Contract Length	fWAR	1.253	1	1.119
	Experience	1.168	1	1.081
	Position	1.140	7	1.009
	Contract Length	1.430	1	1.196
Salary = bWAR + Experience + Position + Contract Length + Team Value	bWAR	1.384	1	1.176
	Experience	1.168	1	1.081
	Position	1.216	7	1.014
	Contract Length	1.446	1	1.202
	Team Value	1.021	1	1.010
Salary = fWAR + Experience + Position + Contract Length + Team Value	fWAR	1.261	1	1.123
	Experience	1.168	1	1.081

	Position	1.149	7	1.010
	Contract Length	1.432	1	1.196
	Team Value	1.018	1	1.009
Salary = WS + Experience + Position + Contract Length + Team Value	WS	1.398	1	1.182
	Experience	1.169	1	1.081
	Position	1.348	7	1.022
	Contract Length	1.342	1	1.158
	Team Value	1.018	1	1.009
Salary = WS CP Adj + Experience + Position + Contract Length + Team Value	WS CP Adj	1.386	1	1.177
	Experience	1.170	1	1.082
	Position	1.328	7	1.020
	Contract Length	1.351	1	1.162
	Team Value	1.019	1	1.009

Table 3. MLB Player Salaries on WAR (in M USD)

Variable	Spotrac		Baseball Reference	
Intercept	7.864***	7.251***	7.408***	6.828***
	(.284)	(.305)	(.256)	(.275)
bWAR	1.451***		1.491***	
	(.146)		(.131)	
fWAR		1.628***		1.633***
		(.148)		(.133)
n	1,265	1,265	1,265	1,265
Adjusted R^2	.072	.087	.092	.106
Residual SE	8.836	8.765	7.949	7.889

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
*, **, *** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 4. MLB Player Salaries on WAR, Experience, and Interaction (in M USD)

Variable	Spotrac				Baseball Reference			
Intercept	- .901*	- 1.761***	.119	- .351	- .557	- 1.357**	- .418	-.181
	(.474)	(.482)	(.518)	(.547)	(.425)	(.432)	(.463)	(.491)
bWAR	1.779***		.706***		1.789***		.763***	
	(.126)		(.260)		(.113)		(.233)	
fWAR		1.922***		.792***		1.963***		.963***
		(.127)		(.259)		(.114)		(.233)
Experience	1.277***	1.293***	1.137***	1.095***	1.159***	1.174***	1.026***	1.009***
	(.059)	(.058)	(.066)	(.069)	(.053)	(.052)	(.059)	(.062)
bWAR × Experience			1.166***				1.159***	
			(.035)				(.032)	
fWAR × Experience				1.187***				1.156***
				(.035)				(.032)
n	1,265	1,265	1,265	1,265	1,265	1,265	1,265	1,265
Adjusted R^2	.321	.342	.334	.357	.341	.360	.354	.372
Residual SE	7.558	7.439	7.496	7.361	6.774	6.672	6.710	6.612

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
*, **, *** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 5. MLB Player Salaries on Win Shares, Experience, and Interaction (in M USD)

Variable	Spotrac						Baseball Reference					
Intercept	6.638***	6.499***	- 1.671**	- 1.686**	- .272	- .180	6.204***	6.069***	- 1.305**	- 1.323**	- .048	.038
	(.351)	(.352)	(.507)	(.504)	(.627)	(.634)	(.316)	(.317)	(.456)	(.454)	(.566)	(.570)
Win Shares	.308***		.337***		.170**		.310***		.336***		.186**	
	(.029)		(.026)		(.052)		(.027)		(.023)		(.046)	
Win Shares (Claim Points)		.328***		.350***		.171**		.330***		.350***		.187***
		(.030)		(.026)		(.053)		(.027)		(.024)		(.048)
Experience			1.217***	1.208***	1.011***	.980***			1.100***	1.090***	.915***	.885***
			(.059)	(.059)	(.081)	(.083)			(.053)	(.053)	(.073)	(.075)
WS × Experience					.025***						.023***	
					(.007)						(.006)	
WS (CP) × Experience						.028***						.025***
						(.007)						(.006)
n	1,265	1,265	1,265	1,265	1,265	1,265	1,265	1,265	1,265	1,265	1,265	1,265
Adjusted R^2	.079	.085	.308	.311	.315	.319	.097	.105	.323	.327	.330	.335
Residual SE	8.803	8.772	7.629	7.612	7.590	7.569	7.929	7.895	6.863	6.844	6.829	6.805

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.

*, **, *** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 6. MLB Player Salaries on WAR, Experience, Position, and Contract Length (in M USD)

Variable (Reference = P)		Spotrac				Baseball Reference			
Intercept		- .995	- 2.186	- 5.765	- 6.371	- .685	- 1.878	- 5.585	- 6.178
		(.503)	(.513)	(.588)	(.578)	(.450)	(.459)	(.515)	(.505)
bWAR		1.721***		.967***		1.721***		.953**	
		(.132)		(.135)		(.118)		(.118)	
fWAR			1.898***		1.215***		1.870***		1.169***
			(.127)		(.130)		(.114)		(.114)
Experience		1.260***	1.268***	1.496***	1.494***	1.140***	1.146***	1.381***	1.379***
		(.060)	(.059)	(.059)	(.058)	(.054)	(.053)	(.051)	(.051)
Position	**1B**	1.511	2.648**	1.345	1.951*	1.586	2.741***	1.417	2.025**
		(.953)	(.922)	(.890)	(.867)	(.853)	(.800)	(.779)	(.758)
	2B	- 3.389***	- 2.037*	- 3.770***	- 3.003***	- 3.025***	- 1.660*	- 3.413***	- 2.653***
		(.925)	(.895)	(.843)	(.843)	(.827)	(.800)	(.756)	(.737)
	3B	2.854**	3.576***	1.964*	2.348**	2.073**	2.817***	1.166	1.556*
		(.897)	(.870)	(.840)	(.822)	(.802)	(.777)	(.735)	(.718)
	C	- 1.054	- .054	- .742	- .150	- 1.080	- .087	- .763	- .185
		(.939)	(.922)	(.878)	(.865)	(.840)	(.824)	(.768)	(.765)
	OF	.645	1.479**	- .099	.387	1.091*	1.936***	.334	.815
		(.587)	(.566)	(.551)	(.538)	(.525)	(.506)	(.482)	(.470)
	SS	.651	1.564	- .771	- .229	.953	1.884*	- .496	.042
		(.997)	(.970)	(.938)	(.920)	(.892)	(.866)	(.821)	(.804)
	DH	1.222	2.299	.431	1.096	2.030*	3.109**	1.224	1.873*
		(1.091)	(1.067)	(1.021)	(1.006)	(.976)	(.954)	(.893)	(.879)
Contract Length				1.357***	1.284***			1.382***	1.319***
				(.100)	(.098)			(.088)	(.086)
n		1,265	1,265	1,265	1,265	1,265	1,265	1,265	1,265

Adjusted R^2	.335	.359	.420	.435	.357	.381	.463	.479
Residual SE	7.478	7.344	6.986	6.892	6.691	6.562	6.114	6.002

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.

*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 7. MLB Player Salaries on Win Shares, Experience, Position, and Contract Length (in M USD)

Variable (Reference = P)		Spotrac		Baseball Reference	
Intercept		- 6.507***	- 6.482***	- 6.302***	- 6.280***
		(.586)	(.586)	(.514)	(.513)
bWAR		.208***		.197***	
		(.028)		(.024)	
fWAR			.216***		.205***
			(.028)		(.025)
Experience		1.491***	1.482***	1.376***	1.367***
		(.059)	(.059)	(.051)	(.051)
Position	**1B**	.337	.358	.506	.520
		(.918)	(.915)	(.805)	(.802)
	2B	- 4.223***	- 4.175***	- 3.804***	- 3.762***
		(.872)	(.869)	(.765)	(.762)
	3B	1.802**	1.903**	1.048	1.141
		(.841)	(.838)	(.737)	(.734)
	C	- .901	- .893	- .899	- .894
		(.877)	(.876)	(.769)	(.768)
	OF	- .567	- .499	- .086	- .024
		(.561)	(.558)	(.492)	(.489)
	SS	- 1.048	- 1.095	- .731	- .778
		(.939)	(.939)	(.824)	(.823)
	DH	- .592	- .554	.272	.304
		(1.034)	(1.032)	(.907)	(.904)
Contract Length		1.417***	1.405***	1.451***	1.439***
		(.096)	(.096)	(.084)	(.084)
n		1,265	1,265	1,265	1,265

Adjusted R^2	.422	.423	.464	.465
Residual SE	6.971	6.964	6.111	6.103

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 8. MLB Player Salaries on Performance (WAR or WS) and Team Valuation (in M USD)

Variable	Spotrac							
Intercept	5.632***	5.031***	4.348***	4.255***	- 3.261***	- 4.117***	- 4.172***	- 4.148***
	(.499)	(.506)	(.532)	(.531)	(.586)	(.588)	(.612)	(.609)
bWAR	1.366***				1.691***			
	(.147)				(.124)			
fWAR		1.551***				1.912***		
		(.147)				(.125)		
Win Shares			1.416***				.323***	
			(.131)				(.025)	
Win Shares (Claim Points)				1.564***				.335***
				(.132)				(.026)
Team Value	1.189***	1.191***	1.235***	1.216***	1.240***	1.245***	1.318***	1.302***
	(.220)	(.218)	(.218)	(.218)	(.187)	(.184)	(.188)	(.188)
Experience					1.281***	1.298***	1.226***	1.216***
					(.058)	(.057)	(.058)	(.058)
n	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263
Adjusted R^2	.092	.107	.101	.107	.344	.366	.335	.337
Residual SE	8.739	8.667	8.696	8.668	7.432	7.310	7.487	7.473

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 8. MLB Player Salaries on Performance (WAR or WS) and Team Valuation (in M USD) - Continued

Variable				**Baseball Reference**				
Intercept	5.424***	4.842***	4.148***	4.057***	- 2.657***	- 3.466***	- 3.552***	- 3.531***
	(.449)	(.455)	(.480)	(.478)	(.526)	(.527)	(.551)	(.548)
bWAR	.295***				1.712***			
	(.029)				(.111)			
fWAR		.314***				1.891***		
		(.030)				(.112)		
Win Shares			.298***				.323***	
			(.026)				(.023)	
Win Shares (Claim Points)				.317***				.336***
				(.027)				(.023)
Team Value	1.057***	1.065***	1.109***	1.089***	1.103***	1.114***	1.184***	1.167***
	(.198)	(.196)	(.196)	(.196)	(.168)	(.165)	(.169)	(.169)
Experience					1.164***	1.179***	1.108***	1.098***
					(.052)	(.051)	(.052)	(.052)
n	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263
Adjusted R^2	.112	.126	.118	.125	.363	.384	.348	.351
Residual SE	7.864	7.801	7.834	7.803	6.663	6.557	6.736	6.720

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.

, *, **** indicates significance at 10%, 5% and 1% level, respectively.*

Table 9. MLB Log Player Salaries on Performance (WAR or WS), Experience, and Interaction (in M USD)

Variable						Spotrac						
Intercept	15.341***	15.255***	15.161***	15.141***	14.173***	14.051***	14.046***	14.043***	14.175***	14.041***	14.002***	14.001***
	(.035)	(.038)	(.043)	(.043)	(.057)	(.058)	(.060)	(.060)	(.063)	(.066)	(.075)	(.076)
bWAR	.174***				.218***				.212***			
	(.018)				(.015)				(.032)			
fWAR		.106***				.254***				.263***		
		(.018)				(.015)				(.031)		
WS			.041***				.045***				.050***	
			(.004)				(.003)				(.006)	
WS (CP)				.044***				.047***				.052***
				(.004)				(.003)				(.006)
Exp					.170***	.173***	.163***	.162***	.169***	.174***	.170***	.168***
					(.007)	(.007)	(.007)	(.007)	(.008)	(.008)	(.001)	(.001)
bWAR × EXP									.001			
									(.004)			
fWAR × EXP										- .001		
										(.004)		
WS × EXP											- .001	
											(.001)	
WS (CP) × EXP												- .001
												(.001)
n	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263
Adjusted R^2	.068	.091	.092	.100	.357	.389	.361	.365	.356	.388	.361	.365
Residual SE	1.096	1.082	1.082	1.077	.911	.887	.908	.905	.911	.888	.908	.905

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 9. MLB Log Player Salaries on Performance (WAR or WS), Experience, and Interaction (in M USD) - Continued

Variable	**Baseball Reference**											
Intercept	15.307***	15.227***	15.130***	15.111***	14.183***	14.070***	14.059***	14.056***	14.197***	14.068***	14.032***	14.02***
	(.035)	(.037)	(.043)	(.043)	(.057)	(.058)	(.060)	(.060)	(.063)	(.066)	(.075)	(.076)
bWAR	.178***				.220***				.205***			
	(.018)				(.015)				(.032)			
fWAR		.204***				.252***				.252***		
		(.018)				(.015)				(.031)		
WS			.041***				.045***				.048***	
			(.004)				(.003)				(.006)	
WS (CP)				.044***				.047***				.050***
				(.004)				(.003)				(.006)
Exp					.164***	.166***	.157***	.156***	.162***	.166***	.161***	.159***
					(.007)	(.007)	(.007)	(.007)	(.004)	(.004)	(.001)	(.001)
bWAR × EXP									.002			
									(.004)			
fWAR × EXP										- .0003		
										(.004)		
WS × EXP											- .0005	
											(.001)	
WS (CP) × EXP												- .0005
												(.001)
n	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263
Adjusted R^2	.072	.091	.094	.102	.346	.372	.349	.352	.345	.372	.348	.351
Residual SE	1.081	1.070	1.069	1.064	.908	.890	.907	.904	.908	.890	.907	.904

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 10. MLB Log Salaries on Performance (WAR or WS), Team Valuation, and Experience (in M USD)

Variable (Reference = P)		Spotrac				Baseball Reference			
Intercept		13.759***	13.664***	13.643***	13.746***	13.652***	13.652***	13.627***	13.631***
bWAR		.150***				.150***			
		(.017)				(.017)			
fWAR			.193***				.187***		
			(.016)				(.016)		
Win Shares				.035***				.034***	
				(.003)				(.003)	
Win Shares Claim Points					.037***				.036***
					(.003)				(.003)
Experience		.188***	.187***	.187***	.185***	.183***	.182***	.182***	.180***
		(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)
Position	**1B**	.132	.125	- .155	- .154	- .027	.068	- .205	- .204
		(.112)	(.108)	(.114)	(.114)	(.111)	(.107)	(.114)	(.113)
	2B	- .377***	- .259***	- .468***	- .461***	- .350***	- .231***	- .435***	- .429***
		(.108)	(.105)	(.109)	(.108)	(.108)	(.104)	(.108)	(.108)
	3B	.242*	.299**	.201	.217*	.191	.251*	.155	.171
		(.105)	(.102)	(.105)	(.104)	(.105)	(.102)	(.104)	(.104)
	C	- .040	- .053	- .072	- .071	- .031	.061	- .060	- .060
		(.110)	(.107)	(.109)	(.109)	(.110)	(.107)	(.109)	(.108)
	OF	.095	.170*	.007	.017	.196**	.272***	.113	.123
		(.069)	(.067)	(.070)	(.069)	(.069)	(.067)	(.070)	(.069)
	SS	- .111	- .027	- .168	- .177	- .028	- .056	- .081	- .091
		(.118)	(.114)	(.117)	(.117)	(.117)	(.114)	(.116)	(.116)
	DH	.139	.243	- .041	- .036	.177	.280*	.004	.009

	(.128)	(.125)	(.129)	(.128)	(.127)	(.125)	(.128)	(.128)
Contract Length	.111***	.098***	.117***	.114***	.110***	.099***	.117***	.115***
	(.013)	(.012)	(.012)	(.012)	(.012)	(.012)	(.012)	(.012)
n	1,255	1,255	1,255	1,255	1,255	1,255	1,255	1,255
Adjusted R^2	.404	.431	.416	.419	.397	.421	.407	.410
Residual SE	.877	.856	.868	.866	.872	.854	.864	.862

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 11. MLB Log Salaries on Performance (WAR or WS), Team Valuation, and Experience (in M USD)

Variable				Spotrac				
Intercept	15.080***	14.998***	14.897***	14.883***	13.896***	13.777***	13.755***	13.757***
	(.062)	(.063)	(.066)	(.065)	(.071)	(.070)	(.073)	(.072)
bWAR	.164***				.208***			
	(.018)				(.015)			
fWAR		.197***				.245***		
		(.018)				(.015)		
WS			.039***				.043***	
			(.004)				(.003)	
WS CP				.042***				.045***
				(.004)				(.003)
Team Value	.139***	.138***	.142***	.140***	.146***	.145***	.153***	.151***
	(.027)	(.027)	(.027)	(.027)	(.007)	(.007)	(.007)	(.007)
Experience					.171***	.173***	.164***	.163***
					(.007)	(.007)	(.007)	(.007)
n	1,263	1,263	1,263	1,263	1,262	1,262	1,262	1,262
Adjusted R^2	.086	.109	.110	.119	.377	.410	.384	.388
Residual SE	1.085	1.072	1.071	1.066	.896	.873	.892	.889

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
, *, **** indicates significance at 10%, 5% and 1% level, respectively.*

Table 11. MLB Log Salaries on Performance (WAR or WS), Team Valuation, and Experience (in M USD) - Continued

Variable	Baseball Reference							
Intercept	15.072***	14.994***	14.890***	14.877***	13.932***	13.820***	13.794***	13.796***
	(.061)	(.062)	(.065)	(.065)	(.071)	(.071)	(.073)	(.073)
bWAR	.169***				.210***			
	(.018)				(.015)			
fWAR		.195***				.242***		
		(.018)				(.015)		
WS			.040***				.043***	
			(.004)				(.003)	
WS CP				.042***				.045***
				(.004)				(.003)
Team Value	.125***	.125***	.129***	.127***	.132***	.132***	.140***	.156***
	(.027)	(.027)	(.027)	(.026)	(.007)	(.007)	(.007)	(.007)
Experience					.164***	.167***	.158***	.156***
					(.007)	(.007)	(.007)	(.007)
n	1,263	1,263	1,263	1,263	1,262	1,262	1,262	1,262
Adjusted R^2	.087	.106	.111	.118	.362	.391	.367	.371
Residual SE	1.073	1.062	1.059	1.055	.896	.878	.893	.891

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 12. MLB Log Salaries on Performance (WAR or WS), Team Valuation, and Interaction Effects (in M USD)

Variable	**Spotrac**				**Baseball Reference**			
Intercept	15.138***	15.068***	14.938***	14.916***	15.148***	15.091***	14.954***	14.901***
	(.069)	(.073)	(.085)	(.086)	(.069)	(.073)	(.084)	(.084)
bWAR	.107**				.093**			
	(.036)				(.035)			
fWAR		.139***				.115***		
		(.037)				(.036)		
WS			.035***				.032***	
			(.007)				(.007)	
WS (CP)				.038***				.039***
				(.007)				(.007)
Team Value	.108***	.100**	.122**	.124**	.085**	.073*	.098**	.115**
	(.032)	(.034)	(.038)	(.038)	(.032)	(.034)	(.037)	(.037)
bWAR × Team Value	.027				.036			
	(.015)				(.015)			
fWAR × Team Value		.029				.040*		
		(.016)				(.016)		
WS × Team Value			.002				.003	
			(.003)				(.003)	
WS (CP) × Team Value				.002				.001
				(.003)				(.003)
n	1,263	1,263	1,263	1,263	1,263	1,263	1,263	1,263
Adjusted R^2	.087	.110	.110	.117	.090	.109	.110	.116
Residual SE	1.084	1.071	1.071	1.067	1.071	1.059	1.059	1.055

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.

*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 13. Correlation Matrix of Key Additionnal Variables

Variable	bWAR	fWAR	WS	WS CP Adj	Experience	Contract Length	Team Value	Age	Signing Bonus
bWAR	1.000								
fWAR	.834	1.000							
WS	.799	.811	1.000						
WS CP Adj	.799	.817	.992	1.000					
Experience	- .121	- .129	- .055	- .042	1.000				
Contract Length	.489	.440	- .320	.380	.382	1.000			
Team Value	.109	.097	- .025	.081	.087	.065	1.000		
Age	- .201	- .249	.836	- .149	- .509	- .509	.165	1.000	
Signing Bonus	.290	.251	.191	.306	.044	.044	.217	.128	1.000

Note: WS CP Adj are Win Shares adjusted for negative claim points.

Table 14. Variance Inflation Factors for Additional Models

Model Specification	Variable	VIF / GVIF	Df	Adjusted GVIF
Salary = bWAR + Exp + Pos + CL + Team Value + Age + Signing Bonus	bWAR	1.469	1	1.212
	Experience	3.951	1	1.988
	Position	1.710	7	1.039
	Contract Length	2.118	1	1.455
	Team Value	1.121	1	1.059
	Age	5.230	1	2.287
	Singing Bonus	1.323	1	1.150
Salary = fWAR + Exp + Pos + CL + Team Value + Age + Signing Bonus	fWAR	1.414	1	1.189
	Experience	3.950	1	1.988
	Position	1.629	7	1.035
	Contract Length	2.223	1	1.491
	Team Value	1.126	1	1.061
	Age	5.230	1	2.287
	Signing Bonus	1.318	1	1.148
Salary = WS + Exp + Pos + CL + Team Value + Age + Signing Bonus	WS	1.540	1	1.241
	Experience	3.990	1	1.997
	Position	1.815	7	1.043
	Contract Length	2.147	1	1.465
	Team Value	1.132	1	1.064
	Age	5.284	1	2.299
	Singing Bonus	1.327	1	1.152
Salary = WS CP Adj+ Exp + Pos + CL + Team Value + Age + Signing Bonus	WS CP Adj	1.496	1	1.223
	Experience	3.972	1	1.993
	Position	1.781	7	1.042

Contract Length	2.125	1	1.458
Team Value	1.128	1	1.062
Age	5.255	1	2.292
Signing Bonus	1.323	1	1.150

Table 15. MLB Player Salaries on WAR, Experience, Position, Contract Length, Age and Signing Bonus (in M USD)

Variable (Reference = P)		Spotrac		Baseball Reference	
Intercept		- 16.570*	- 16.850*	- 13.970*	- 14.270*
		(6.656)	(6.614)	(6.562)	(6.525)
bWAR		.225		.284	
		(.226)		(.223)	
fWAR			.517*		.526*
			(.250)		(.246)
Experience		1.331***	1.330***	1.445***	1.444***
		(.217)	(.216)	(.214)	(.213)
Position	**1B**	3.708	3.891	3.695	3.946
		(2.228)	(2.189)	(2.196)	(2.159)
	2B	- 4.510*	- 4.154*	- 4.035*	- 3.607*
		(1.863)	(1.832)	(1.86)	(1.807)
	3B	4.167**	4.236**	4.467**	4.607**
		(1.478)	(1.451)	(1.457)	(1.431)
	C	- 5.693*	- 5.159*	- 6.681*	- 6.133*
		(2.695)	(2.690)	(2.657)	(2.653)
	OF	1.024	1.345	1.424	1.782
		(1.017)	(1.007)	(1.003)	(.994)
	SS	1.974	2.081	2.376	2.578
		(1.680)	(1.631)	(1.656)	(1.609)
	DH	- .427	- .048	- .160	.246
		(2.061)	(2.052)	(2.032)	(2.024)
Contract Length		1.004***	.913***	.995***	.917***
		(.202)	(.205)	(.199)	(.202)
Age		.001	.461	.326	.327

	(.239)	(.238)	(.236)	(.235)
Signing Bonus	.001***	.001***	.002***	.002***
	(.000)	(.000)	(.000)	(.000)
n	268	268	268	268
Adjusted R^2	.500	.506	.524	.530
Residual SE	6.332	6.292	6.242	6.207

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
, **, * indicates significance at 10%, 5% and 1% level, respectively.*

Table 16. MLB Player Salaries on Win Shares, Experience, Position, Contract Length, Age and Signing Bonus (in M USD)

Variable (Reference = P)		Spotrac		Baseball Reference	
Intercept		- 16.370*	- 16.470*	- 13.710*	- 13.840*
		(6.678)	(6.665)	(6.587)	(6.574)
WS		.034		.044	
		(.053)		(.052)	
WS (CP)			.041		.052
			(.054)		(.053)
Experience		1.346***	1.345***	1.464***	1.464***
		(.218)	(.218)	(.215)	(.215)
Position	**1B**	3.644	3.588	3.613	3.543
		(2.287)	(2.282)	(2.256)	(2.250)
	2B	- 4.493*	- 4.510*	- 4.015*	- 4.035*
		(1.885)	(1.877)	(1.859)	(1.851)
	3B	4.238**	4.244**	4.564**	4.573**
		(1.483)	(1.475)	(1.463)	(1.455)
	C	- 5.743*	- 5.749*	- 6.745*	- 6.752*
		(2.700)	(2.699)	(2.664)	(2.662)
	OF	.999	1.000	1.329	1.394
		(1.037)	(1.029)	(1.023)	(1.015)
	SS	2.111	2.074	2.548	2.503
		(1.678)	(1.676)	(1.656)	(1.653)
	DH	- .531	- .546	- .292	- .310
		(2.081)	(2.077)	(2.053)	(2.049)
Contract Length		1.022***	1.017***	1.018***	1.012***
		(.202)	(.202)	(.200)	(.199)
Age		.445	.448	.307	.310

	(.240)	(.240)	(.237)	(.237)
Signing Bonus	.002***	.002***	.002***	.002***
	(.000)	(.000)	(.000)	(.000)
n	268	268	268	268
Adjusted R^2	.500	.506	.524	.530
Residual SE	6.339	6.337	6.253	6.250

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
, **, * indicates significance at 10%, 5% and 1% level, respectively.*

Table 17. MLB Player Salaries on Performance (WAR or WS), Team Valuation, Experience and Signing Bonus (in M USD)

Variable	Spotrac				Baseball Reference			
Intercept	- .588	- 1.794	- 1.274	- 1.253	- 1.266	- 2.451	- 2.022	- 2.011
	(1.216)	(1.257)	(1.298)	(1.297)	(1.210)	(1.255)	(1.294)	(1.293)
bWAR	.773***				.893***			
	(.211)				(.210)			
fWAR		1.114***				1.188***		
		(.235)				(.234)		
Win Shares			.165***				.187***	
			(.049)				(.049)	
Win Shares (Claim Points)				.167***				.191***
				(.050)				(.050)
Team Value	.933*	1.005*	.988*	.971*	.973*	1.051*	1.036*	1.018*
	(.452)	(.445)	(.454)	(.454)	(.450)	(.444)	(.453)	(.453)
Experience	1.338***	1.381***	1.331***	1.329***	1.331***	1.372***	1.321***	1.320***
	(.122)	(.121)	(.122)	(.122)	(.121)	(.120)	(.122)	(.122)
Signing Bonus	.002***	.002***	.002***	.002***	.002***	.002***	.002***	.002***
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
n	268	268	268	268	268	268	268	268
Adjusted R^2	.425	.433	.420	.420	.442	.457	.436	.436
Residual SE	6.788	6.679	6.814	6.816	6.756	6.666	6.798	6.798

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.

, **, * indicates significance at 10%, 5% and 1% level, respectively.*

Table 18. MLB Log Salaries on Performance (WAR or WS), Experience, Position, Contract Length, Age and Signing Bonus (in M USD)

Variable (Reference = P)		Spotrac				Baseball Reference			
Intercept		12.900***	12.850***	12.940***	12.920***	13.550***	13.510***	13.600***	13.580***
		(.692)	(.686)	(.696)	(.694)	(.683)	(.678)	(.687)	(.685)
bWAR		.043				.049*			
		(.024)				(.023)			
fWAR			.072**				.071**		
			(.026)				(.026)		
Win Shares				.007				.008	
				(.006)				(.005)	
Win Shares Claim Points					.008				.009
					.008				(.006)
Experience		.132***	.132***	.135***	.135***	.158***	.158***	.161***	.161***
		(.023)	(.022)	(.023)	(.023)	(.022)	(.022)	(.022)	(.022)
Position	**1B**	.026	.069	.008	.002	.057	.110	.037	.032
		(.232)	(.227)	(.238)	(.238)	(.228)	(.225)	(.235)	(.235)
	2B	- .551**	- .487*	- .551***	- .552***	- .491*	- .420*	- .491*	- .491*
		(.194)	(.190)	(.196)	(.196)	(.191)	(.188)	(.194)	(.193)
	3B	.280	.302*	.290	.294	.342*	.372*	.355*	.359*
		(.154)	(.151)	(.155)	(154)	(.152)	(.149)	(.153)	(.152)
	C	- .819**	- .743**	- .830**	- .830**	- .810**	- .735**	- .822**	- .822**
		(.280)	(.105)	(.281)	(.281)	(.276)	(.276)	(.278)	(.278)
	OF	.121	.172	.114	.116	.191	.245*	.183	.186
		(.106)	(.105)	(.108)	(.107)	(.104)	(.103)	(.107)	(.106)
	SS	- .033	.002	- .010	- .014	.033	.079	.060	.056
		(.175)	(.169)	(.175)	(.175)	(.172)	(.167)	(.173)	(.172)
	DH	.023	.080	.000	- .001	.065	.123	.123	.039

	(.214)	(.213)	(.217)	(.216)	(.211)	(.210)	(.214)	(.214)
Contract Length	.049*	.039	.051*	.051*	.036	.029	.040	.039
	(.021)	(.021)	(.021)	(.021)	(.021)	(.021)	(.021)	(.021)
Age	.062*	.062*	.059*	.059*	.033	.033	.030	.030
	(.021)	(.025)	(.025)	(.021)	(.025)	(.024)	(.025)	(.025)
Signing Bonus	.140**	.131**	.147***	.145***	.152***	.147***	.159***	.158***
	(.043)	(.042)	(.043)	(.043)	(.042)	(.042)	(.042)	(.042)
n	268	268	268	268	268	268	268	268
Adjusted R^2	.505	.513	.502	.503	.529	.534	.524	.525
Residual SE	.658	.653	.660	.660	.649	.645	.652	.652

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.

*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Table 19. MLB Log Salaries on Performance (WAR or WS), Team Valuation, Experience and Signing Bonus (in M USD)

Variable	Spotrac				Baseball Reference			
Intercept	14.530*** (.121)	14.430*** (.126)	14.490*** (.129)	14.490*** (.129)	14.450*** (.119)	14.350*** (.124)	14.400*** (.128)	14.400*** (.127)
bWAR	.054* (.021)				.066** (.021)			
fWAR		.088*** (.023)				.091*** (.023)		
WS			.011* (.005)				.013** (.005)	
WS (CP)				.011* (.005)				.013** (.005)
Team Value	.128** (.045)	.133** (.044)	.131** (.045)	.130** (.045)	.124** (.044)	.130** (.044)	.128** (.045)	.127** (.045)
Experience	.155*** (.012)	.159*** (.012)	.155*** (.012)	.155*** (.012)	.159*** (.012)	.162*** (.012)	.158*** (.012)	.158*** (.012)
Signing Bonus	.165*** (.042)	.154*** (.041)	.166*** (.043)	.167*** (.043)	.178*** (.042)	.173*** (.041)	.181*** (.042)	.182*** (.042)
n	268	268	268	268	268	268	268	268
Adjusted R^2	.478	.492	.475	.475	.504	.513	.499	.499
Residual SE	.676	.667	.678	.678	.666	.660	.669	.670

Note: Salaries and coefficients are reported in millions of USD (M USD). Standard errors are reported in parentheses.
*, ***, **** *indicates significance at 10%, 5% and 1% level, respectively.*

Appendix B – Abbreviations and terms

Standard Batting

Abbreviation	**Term**
Pos	Position
C	Catcher
1B	First Baseman
2B	Second Baseman
SS	Shortstop
3B	Third Baseman
LF	Left Fielder
CF	Center Fielder
RF	Right Fielder
DH	Designated Hitter
WAR	Wins Above Replacement for position player
G	Games Played
AB	At Bats
RS	Runs Scored
H	Hits
2B	Doubles
3B	Triples
HR	Home Runs
RBI	Runs Batted In
SB	Stolen Bases
CS	Caught Stealing
BB	Base on Balls/Walks
SO	Strikeouts
GIDP	Double Plays Grounded Into
HBP	Times Hit by a Pitch
SH	Sacrifice Hits
SF	Sacrifice Flies
IBB	Intentional Bases on Balls

Standard Pitching

Abbreviation	**Term**
Pos	Position
SP	Starting Pitcher
RP	Relief Pitcher
CL	Closer
WAR	Wins Above Replacement for Pitchers
W	Wins
L	Losses
G	Games Pitched

GS	Games Started
GF	Games Finished
CG	Complete Games
SHO	Shutouts
SV	Saves
IP	Innings Pitched
H	Hits Allowed
R	Runs Allowed
ER	Earned Runs Allowed
HR	Home Runs Allowed
BB	Bases on Balls/Walks
IBB	Intentional Bases on Balls
SO	Strikeouts
HBP	Times Hit by a Pitch
BK	Balks
WP	Wild Pitches
BF	Batters Faced

Standard Fielding

Abbreviation	**Term**
G	Games Played
GS	Games Started
CG	Complete Games at a Single Position
Inn	Innings Played at Position
PO	Putouts
A	Assists
E	Errors Committed
DP	Double Plays Turned
PB	Passed Balls
WP	Wild Pitches
SB	Stolen Bases
CS	Caught Stealing
Pick	Pickoffs

Appendix C – Metrics and formulas

Standard Batting

Plate Appearances

$$PA = AB + BB + HBP + SF + SH$$

Batting Average

$$BA = \frac{H}{AB}$$

On-Base Percentage

$$OBP = \frac{H + BB + HBP}{AB + BB + HBP + SF}$$

Slugging Percentage

$$SLG = \frac{1B + (2 \times 2B) + (3 \times 3B) + (4 \times HR)}{AB}$$

On-base Plus Slugging

$$OPS = OBP + SLG$$

On-base Plus Slugging Plus

$$OBP = 100 \times \left[\frac{OBP}{lgOBP} + \frac{SLG}{lgSLG} - 1\right]$$

Total Bases

$$TB = 1B + (2 \times 2B) + (3 \times 3B) + (4 \times HR)$$

Standard Pitching

Win-Loss Percentage

$$W - L\% = \frac{W}{(W + L)}$$

Earned Run Average

$$ERA = 9 \times \frac{ER}{IP}$$

Earned Run Average Plus

$$ERA+ = 100 \times \left[\frac{lgERA}{ERA}\right]$$

Fielding Independent Pitching

$$FIP = \frac{(13 \times HR) + [3 \times (BB + HBP)] - (2 \times K)}{IP} + Constant$$

Walks plus Hits per Inning Pitched

$$WHIP = \frac{BB + H}{IP}$$

Hit per 9 Innings

$$H9 = 9 \times \frac{H}{IP}$$

Home Runs per 9 Innings

$$HR9 = 9 \times \frac{HR}{IP}$$

Walks per 9 Innings

$$BB9 = 9 \times \frac{BB}{IP}$$

Strikeouts per 9 Innings

$$SO9 = 9 \times \frac{SO}{IP}$$

Strikeout-to-Walk Ratio

$$SO/BB = 100 \times \frac{SO}{BB}$$

Standard Fielding

Defensive Chances

$$Ch = PO + A + E$$

Fielding Percentage

$$Fld\% = \frac{PO + A}{PO + A + E}$$

Range Factor per 9 Innings

$$RF/9 = 9 \times \frac{PO + A}{IP}$$

Range Factor per Games Played

$$RF/G = 9 \times \frac{PO + A}{GP}$$

Caught Stealing Percentage

$$CS\% = \frac{CS}{SB + CS}$$

Appendix D – Park Factors

This appendix presents the procedure used in Bill James's *Win Shares* system to compute park factors. A park factor is a multiplier that adjusts team statistics for the run environment of the home ballpark.

i. Definitions and notations

Let,

- RS_{home}: runs scored by the team in home games
- RA_{home}: runs allowed by the team in home games
- RS_{road}: runs scored by the team in road games
- RA_{road}: runs allowed by the team in road games
- G_{home}: number of home games
- G_{road}: number of road games

ii. Methodology

Compute total runs in a home game

$$R_{home} = RS_{home} + RA_{home}$$

Computer total runs in a road game

$$R_{road} = RS_{road} + RA_{road}$$

Convert to per-game value

$$R/G_{home} = \frac{R_{home}}{G_{home}}$$

$$R/G_{road} = \frac{R_{road}}{G_{road}}$$

Compute the raw park factor

$$Raw\ Park\ Factor = \frac{R/G_{home}}{R/G_{road}}$$

Apply the half-impact adjustment

$$Adjusted\ Park\ Factor = 1 + \frac{Raw\ Park\ Factor - 1}{2}$$

Convert to a park multiplier

$$Park\ Multiplier = \frac{1}{Adjusted\ Park\ Factor}$$

Adjust team statistics

$$Park\ Adjusted\ Runs\ Created = RC \times Park\ Multiplier$$

$$Park\ Adjusted\ Runs\ Allowed = RA \times Park\ Multiplier$$

Appendix D – Calculating Bill James' Win Shares

This appendix presents a reconstruction of the Win Shares system developed by Bill James (2002). A Win Share is an allocation of team wins to individual players based on offensive, pitching, and fielding contributions.

i. System Structure

Accounting identity

Each team win corresponds to three Win Shares.

$$\sum_{i \in t} WS_i = 3 \times W_t$$

Component decomposition

Each player's total Win Shares are the sum of offensive, pitching and fielding components.

$$WS_i = WS_{off,i} + WS_{pit,i} + WS_{fld,i}$$

ii. League Context

League run environment

League scoring level used as baseline for marginal runs.

Let

- RS_t: runs scored by team t
- G_t: games played by team t
- N_{teams}: number of teams in the league

League average runs per team

$$LG = \frac{\sum_t RS_t}{N_{teams}}$$

League runs per game

$$R/G_{lg} = \frac{\sum_t RS_t}{\sum_t G_t}$$

League defensive efficiency

Measures how effectively the league converts ball into plays into outs.

Let

- PO_t: putouts recorded by team t
- SO_t: strikeouts recorded by team t
- BIP_t: balls in play against team t

- AB_t: at-bats against team t
- HR_t: home runs allowed by team t
- SF_t: sacrifice flies allowed by team t

Balls in play

$$BIP_t = (AB_t + SO_t + HR_t) + SF_t$$

League defensive efficiency ratio

$$DER_{lg} = \frac{\sum_t (PO_t - SO_t)}{\sum_t BIP_t}$$

iii. Team Marginal Runs and Win Shares

Marginal offensive runs
Teams runs above a league-defined baseline.

Let

- $\overline{RS}$: league average runs scored per team

$$Off_t = RS_t - \frac{\overline{RS}}{2}$$

Marginal defensive runs
Team runs prevented relative to a league baseline.

Let

- RA_t: runs prevented by team t

$$Def_t = \frac{\overline{RS}}{2} - RA_t$$

Expected win percentage
Team performance expressed relative to league scoring conditions.

$$Win\%_t = \frac{Off_t + Def_t}{2\overline{RS}}$$

Team Win Shares
Total Win Shares allocated to the roster.

$$WS_{team,t} = 3W_t$$

iv. Offensive vs. Defensive Allocation

Let

- $f_{off,t}$: offensive allocation fraction of team t
- $f_{def,t}$: defensive allocation fraction of team t
- $WS_{off,t}$: team t offensive win shares
- $WS_{def,t}$: team t defensive win shares

$$f_{off,t} = \frac{Off_t}{Off_t + Def_t}$$

$$f_{def,t} = \frac{Def_t}{Off_t + Def_t}$$

$$WS_{off,t} = WS_t \times f_{off,t}$$

$$WS_{def,t} = WS_t \times f_{def,t}$$

v. Pitching vs. Fielding Allocation

Let

- H_t: hits by team t
- BB_t: walks by team t
- IP_t: innings pitched by team t

Component pitching runs
Estimated pitching responsibility for runs allowed.

Let

- $cERA_t$: component earned run average by team t
- CRA_t: component runs allowed by team t

$$cERA_t = \frac{2.13 \times HR_t + 0.47 \times BB_t - 0.34 \times SO_t + 0.37 \times H_t}{IP_t} \times 9$$

$$CRA_t = \frac{cERA_t}{9} \times IP_t$$

Fielding-dependant runs

$$FDR_t = RA_t - CRA_t$$

Allocation fractions

Let

- $f_{pit,t}$: pitching allocation fraction of team t
- $f_{fdl,t}$: fielding allocation fraction of team t

- $WS_{pit,t}$: team t offensive win shares
- $WS_{fld,t}$: team t defensive win shares

$$WS_{pit,t} = WS_{def,t} \times f_{pit,t}$$

$$WS_{fld,t} = WS_{def,t} \times f_{fdl,t}$$

vi. Park Factor Adjustments

Let

- PF_t: team t park factor

Park Factor

$$PF_t = \frac{\frac{RS_{home,t} + RA_{home,t}}{G_{home,t}}}{\frac{RS_{road,t} + RA_{road,t}}{G_{road,t}}}$$

Adjusted statistics

Let

- $RC_{adj,b}$: park adjusted runs created of hitter b
- $RA_{adj,p}$: park adjusted runs allowed of pitcher p

$$RC_{adj,i} = \frac{RC_i}{PF_t}$$

$$RA_{adj,i} = RA_i \times PF_t$$

vii. Offensive Win Shares

Let

- TB_b: total bases of batter b
- SB_b: stolen bases of batter b
- CS_b: caught stealing of batter b
- HBP_b: hit by pitches of batter b
- $GIDP_b$: grounded into double of batter b
- SH_b: sacrifice hits of batter b
- IBB_b: intentional base on balls of batter b
- MR_b: marginal runs of batter b
- k_t: scaling factor of team t

Runs created

Let

- $A_b = H_b + BB_b - CS_b + HBP_b - GDIP_b$
- $B_b = TB_b + 0.26 \times (BB_b - IBB_b + HBP_b) + 0.52 \times (SH_b + SF_b + SB_b)$
- $C_b = AB_b + BB_b + HBP_b + SH_b + SF_b$

$$RC_b = \frac{A_b \times B_b}{C_b}$$

Baseline
A fixed 52% of Runs Created is treated as baseline.

$$RC_{base,b} = 0.52 \times RC_{adj,b}$$

Marginal Runs

$$MR_{off,b} = RC_{adj,b} - RC_{base,b}$$

Scaling

$$k_{off,t} = \frac{Off_t}{\sum_b MR_{off,b}}$$

$$MR_{off,b}^{scaled} = MR_{off,b} \times k_{off,b}$$

Offensive Win Shares

$$WS_{off,b} = WS_{off,t} \times \frac{MR_{off,b}^{scaled}}{Off_t}$$

viii. Pitching Win Shares

Let

- MR_p: marginal runs of pitcher p
- IP_p: innings pitched by pitcher p

Pitching Marginal Runs

$$MR_p = RA_{adj,p} - \frac{CRA_t}{IP_t} \times IP_p$$

Scaling

$$k_{pit,t} = \frac{CRA_t}{\sum_p MR_p}$$

$$MR_p^{scaled} = MR_p \times k_{pit,t}$$

Pitching Win Shares

$$WS_p = WS_{pit,t} \times \frac{MR_p^{scaled}}{CRA_t}$$

ix. Pitching Win Shares

Let

- FDR_t: team t fielding-dependent runs
- Inn_f: defensive innings played by fielder f
- A_f: assists of fielder f
- E_f: errors of fielder f
- $WS_{fld,t}$: team t fielding win shares
- $MR_{fld,f}$: marginal fielding runs by field f

Fielding marginal runs

$$MR_{fld,f} = PO_f + A_f - E_f$$

Scaling

$$k_{fld,t} = \frac{FDR_t}{\sum_f MR_{fld,f}}$$

$$MR_{fld,f}^{scaled} = MR_{fld,f} \times k_{fld,t}$$

Fielding Win Shares

$$WS_{fld,f} = WS_{fld,t} \times \frac{MR_{fld,f}^{scaled}}{FDR_t}$$

x. Positional Adjustments

Let

- PA_i: plate appearances at position i
- $Pos_{adj,i}$: positional adjustment run value for position i
- $Pos_{run,i}$: positional runs for position i
- $WS_{pos,t}$: total positional adjustment Win Shares

Positional run value
Each defensive position receives a fixed positional run value.

$$Pos_{adj,i} \in \{1B, 2B, 3B, SS, LF, CF, RF, C\}$$

$$C: +1$$

$$SS: +7$$

$$2B: +3$$

$$3B: +2$$

$$CF: +2$$

$$RF: -7$$

$$LF: -7$$

$$1B: -9$$

Player positional adjustment

$$Pos_{run,i} = Pos_{adj,i} \times \frac{Inn_i}{1458}$$

$$1458 = team\ defensive\ innings$$

Scaling

$$k_{pos,i} = \frac{\sum_i Pos_{run,i}}{\sum_i MR_{fld,i}^{scaled}}$$

$$Pos_{run,i}^{scaled} = Pos_{run,i} \times k_{pos,i}$$

Positional Win Shares

$$WS_{pos,i} = WS_{fld,t} \times \frac{Pos_{run,i}^{scaled}}{\sum_i Pos_{run,i}}$$

xi. Final Player Win Shares

Let

- $WS_{off,b}$: offensive win shares of batter b
- $WS_{pit,p}$: pitching win shares of pitcher p
- $WS_{fld,f}$: fielding win shares of fielder f
- $WS_{pos,i}$: positional win shares of player i

Total Win Shares

Batters

$$WS_b = WS_{off,b} + WS_{fld,b} + WS_{pos,b}$$

Pitchers

$$WS_p = WS_{pit,p}$$

Team Identity

$$\sum_b WS_b + \sum_p WS_p = WS_t$$

$$WS_t = 3 \times W_t$$

Appendix E – Calculating Claim Points

Claim Points quantify each player's marginal contribution to offense, defense, and pitching relative to context-adjusted baselines. Team Win Shares for each component are allocated in proportion to Claim Points.

i. Offensive Claim Points

Let

- RC_i: runs created by player i
- $(RC/O)_{Lg}$: league runs created per out
- O_i: outs made by player i
- $WS_{off,t}$: team offensive win shares
- $RC_{base,i}$: player i baseline runs
- $CP_{off,i}$: player i offensive claim points

Baseline Runs

Number of runs a player is expected to create based on league-average production per out.

$$RC_{base,i} = 0.52 \times (RC/O)_{Lg} \times O_i$$

Offensive Claim Points

Measures a player's runs created above the baseline level of expected production.

$$CP_{off,i} = RC_i - RC_{base,i}$$

Team Offensive Claim Points

$$CP_{off,t} = \sum_i CP_{off,i}$$

Offensive Win Shares allocation

$$WS_{off,i} = WS_{off,t} \times \frac{CP_{off,i}}{CP_{off,t}}$$

ii. Defensive Claim Points

Let

- $E_{i,k}$: defensive plays where $k \epsilon \{PO, A, E, DP, WP, SB, CS\}$
- w_k: weight for type of play
- $\overline{Def_i}$: expected defensive value at position i
- $r_{i,k}$: league-average rate of event k at position i
- $Opp_{i,k}$: player i opportunities for even k
- $F\%_{base}$: baseline fielding percentage

Raw Defensive Value

$$Def_i = \sum_k w_k \times E_{i,k}$$

Marginal Defensive Value

Expected number of weighted defensive events for a league-average player at the same position and playing time.

$$r_{i,k} = \frac{League\ total\ of\ evenk\ k\ at\ position\ i}{League\ total\ opportunities\ at\ position\ i}$$

$$Opp_{i,k} = player\ opportunities\ for\ event\ k$$

where opportunities are measured using:

- innings at position i
- chances (putouts, assists, errors)
- games caught (catcher-specific events)
- outfield innings (outfielder assists and errors)
- first-base chances

Baseline Defensive value

$$\overline{Def_i} = \sum_k w_k \times \left(r_{i,k} \times Opp_{i,k}\right)$$

Defensive Claim Points

$$CP_{i,def} = Def_i - \overline{Def_i}$$

Error-avoidance component (positional)

Credits a player for errors avoided relative to the baseline fielding percentage at position.

$$CP_{i,ErrAv} = [(PO_i + A_i + E_i) \times (1 - F\%_{base})] - E_i$$

Team defensive Claim Points

$$CP_{def,t} = \sum_i CP_{def,i}$$

iii. Pitching Claim Points

Let

- $(RA/IP)_{Lg}$: league runs allowed per innings pitched
- $\overline{RA}_{base,i}$: expected runs allowed for league-average pitching
- $CP_{pit,i}$: player i pitching claim points

Baseline runs allowed

$$\overline{RA}_{base,i} = (RA/IP)_{Lg} \times IP_i$$

Pitching Claim Points

$$CP_{pit,i} = \overline{RA}_{base,i} - RA_i$$

Team Pitching Claim Points

$$CP_{pit,t} = \sum_{i} CP_{pit,i}$$

iv. Truncation Rule

The truncation rule ensures that no player receives negative Win Shares even when their claim points are negative.

$$WS_i = \max(0, WS_i)$$