

The Application and Impact of Numerical Models for Decision-Making in Sports

Jordan James Foley, Roi Guinto, Nicola Ricci, Marcus Wu Shihong

*Technology and Policy Program, Engineering Systems Division
Massachusetts Institute of Technology, Cambridge, MA, 02139*

This paper describes the viability and impact of statistical information and numerical models for decision-making in professional sports with regards to performance assessment and team management. Sports statistics may have established winning formulas and transformed the face of modern sports, but they do not always take precedence in determining decision outcomes. A case study in American football is used, where the decision to keep, cut, or trade renowned quarterback Ben Roethlisberger by the Pittsburgh Steelers was not only influenced by his stellar performances and team contributions, but also by external pressures and his off-field behavior.

I. Introduction

The advent of sports statistics is fast changing the way professional sports are being played in the era of free information and cutting-edge computation. Beginning with sports enthusiasts and fantasy league players who gather and analyze information to predict results, sports statistics have rapidly evolved in recent years with the development of various tools, techniques and metrics to better measure both player and team performance. The usage of such quantitative methods, instead of reliance on qualitative human expertise and intuition, has greatly transformed team management and strategy development in modern-day sports. Early adopters such as baseball's Oakland Athletics, basketball's Seattle Supersonics and soccer's AC Milan¹ have leveraged statistical analysis to calculate their way towards success. As such, data collection and the application of statistical methods are now ubiquitous in every major sporting arena for decision-making.

II. Evolution and Application of Sports Statistics

Sports statistics of the past were mere collections of crude data articulated by domain experts like coaches, managers and owners, who make decisions on various time scales. Coaches consider the day-to-day management of their players; general managers handled contracts operating on time scales that can extend to a decade; and owners considered their long-term investment strategies. The absence of synchronization and practical statistical methods for rational decision-making posed numerous risks. Misguided decisions in the front office can potentially upset the financial status of a team for several years. Wrong calls by a coach, based on imprecise statistics or intuition, can spell defeat in a single game and even an entire season. These choices often weighed the merits of statistical analysis against conventional wisdom, leading to initial reluctance in adopting the former. A recent example in American football is the decision made by Bill Belichick, coach of the New England Patriots, to pass on the fourth down in the 2009 playoffs². Despite being statistically sound, the decision backfired and led to a debate as to whether the correct statistics were used, or whether the statistics would ever be able to fully account for situational details.

Methods for assessing performance have existed since the inception of professional sports. Players in sports leagues have been traditionally analyzed by a combination of objective and subjective information. For example, baseball scouts use a list known as the Five Tools, which include hitting for power, hitting for average, speed, arm strength, and defensive range. In the past, traits such as speed and defensive range could not be quantified easily and required the subjective judgment of scouts. These scouts were often tied to the game as former players or lifelong assistants and coaches and developed an intuition on these statistics³. Information also came from analyzing intangibles, which could never be quantified. According to John Sanders, there are traits such as work ethic and stress management, which are vital to the continued improvement of a player. In-house statisticians were then hired to provide measures of performance and decision-making criteria. Clearly, statistics were beginning to play a crucial role in decision-making for sports to provide an objective analysis of a player's past performance.²

The rigorous use of statistics for decision-making in sports began with Bill James, who published "Bill James Baseball Abstracts"⁴ annually to criticize traditional baseball performance metrics that were often

imprecise, and publicize his compilation of radical ranking formulae and statistical performance measures more famously known as “sabermetrics”. Metrics such as Runs Created (RC), Earned Runs Average (ERA) and Total Player Rating (TPR) became popularized. Basketball later became acquainted with Dean Oliver’s “ABPRmetrics”⁵, named for the Association of Professional Basketball Researchers (ABPR), and relied on shot zone analysis, player efficiency ratings (PER) and adjustments to plus/minus ratings for measuring players’ contributions to victories. Football also grew to become statistically driven with metrics like Adjusted Line Yards (ALY) to allocate credit to an offensive line based on the yardage across which the ball is carried, and the Defense-Adjusted Value Over Average (DVOA) to comparatively measure the success of a particular play. As such, the use of statistics is now prevalent in sports. Recognizing hidden trends in data facilitate decision-making and the development of new strategies that can create arbitrage opportunities for victory.

With the evolution of professional sports as a lucrative trade, data collection grew in scope and magnitude, and sporting franchises required more practical methods of valuable knowledge using increasingly abstract data mining techniques. Data can comprise individual player performance, overall team performance, coaching decisions, managerial decisions, and discrete game-based events. Through performance metrics, and statistical algorithms, data can be translated into actionable knowledge that sporting franchises can apply to gain competitive advantages over their peers. Such a knowledge-driven approach is now applied extensively throughout a franchise. Players can reflect on past performances and improve their game-time performance using video analysis techniques; scouts can accurately identify potential talents in sports draft using statistical analysis and projection techniques; coaches and managers can analyze player performances to manage team selections and use machine learning and simulation techniques to find optimal tactics; owners can study overall team performances over years to determine profitable investment and marketing strategies. Outside the franchise, betting companies invest heavily on sports statistics in setting payouts for score predictions, and sports fans rely on such data to track their favorite players and teams or forecasting exact scores.

However, sports statistics also present new obstacles⁶. Performance metrics are often misused and imprecise, as they do not measure what was intended. Two of many such examples are the frequent use of number of receptions and yards per carry for all football players. The number of receptions is the number of times a player catches a forward pass while yards per carry measures how far a player can carry the ball forward. However, they often lead to unfair comparisons as different playing styles heavily influence such metrics. More emphasis should instead be placed on the points scored. Quantifying risk is also another problem in sports statistics. For example, a defensive back in football can either take the risk by attempting to intercept the opponent’s ball or play safe through solid cover defense. A player can thus have greater perceived player value if he takes the risks and performs successful interceptions often. Such unobvious yet valuable attributes make the assessing of player’s performance difficult. As such, these drawbacks often lead to high inertia in their initial usage and may further complicate decision-making.

Despite their inherent uncertainties, sports statistics remain crucial in translating interesting patterns in data to sound decisions that positively transform the fortunes of a sporting franchise. With monetary investments, reputation and fanbases at stake, franchises will be determined to avoid making bad decisions based on the absence or improper usage of statistics. Sports statistics will continue to evolve to replace traditional, imprecise measures with mathematically rigorous and salient ones, and their credibility will be reaffirmed with increasing application and frequency of success. Therefore, statistics will continue transforming the ways in which decisions are made in sports.

III. American Football Case Study: Ben Roethlisberger

Sports statistics may have established winning formulas and transformed the face of modern sports, but they do not always take precedence in determining decision outcomes. A case study in American football is used, where the decision to keep, cut, or trade renowned quarterback Ben Roethlisberger at the Pittsburgh Steelers was not only influenced by his stellar performances or team contributions, but also by external pressures and his off-field behavior. An overview of the sport, an analysis of the numerical models and the player’s history will be described while the decision-making process will be discussed in detail.

1. An Overview of American Football and the Quarterback

American Football is a sport where teams field 11 players at a time in order to move a ball to the other team's end zone. Each player specializes in a certain role such as quarterback or tight end and plays on either offense or defense. By either throwing or running with the ball, the offense attempts to move down the field before the ball carrier is tackled by a defensive player. The quarterback is crucial to the offense since he is responsible for either throwing the ball to a receiver or handing the ball off to a receiver. He is involved in nearly every play of the offense and functions as a de facto team leader.

2. Numerical Models: Passer Rating and QBR

Statistics-based metrics have widely been used over the past forty years to characterize the performance of quarterbacks (QB). A variety of metrics have been developed, but two preeminent ones are considered and described here: Passer rating (PR) and Total Quarterback Rating (QBR).

Passer Rating: Devised in 1971, PR is the most famous and used metric to rate quarterbacks. Prior to its inception, the NFL had struggled for years on how to devise a metric that would efficiently and realistically assess a QB's passing ability. The metric takes into account four different factors (categories): Completion Percentage, Average Yards Per Attempt, Percentage of Touchdown Passes, and Percentage of Interceptions. These are represented below by the metrics 'a', 'b', 'c' and 'd' respectively. For any of these categories, it is impossible to score lower than 0 or higher than 2.735: which is, if the QB's score exceeds 2.735 or is lower than 0, it is automatically set to 2.735 and 0, respectively. In order to obtain the highest possible score of 158.3, a quarterback must score: (1) at least a 77.5% in Completion Percentage; (2) at least 12.5 in average yards per attempt; (3) at least 11.875% in percentage of touchdown passes; and (4) 0% in Percentage of Interception. The actual formula used to calculate overall PR, as well as the score in each individual category, is available to the public and shown below. The overall Passer Rating is thus a numerical model that aggregates the four metrics⁷.

$$\begin{aligned}
 a &= \left(\frac{COMP}{ATT} - 0.3 \right) \times 5 && \text{Where} \\
 b &= \left(\frac{YARDS}{ATT} - 3 \right) \times 0.25 && \text{ATT= Number of passing attempts} \\
 c &= \left(\frac{TD}{ATT} \right) \times 20 && \text{COMP = Number of completions} \\
 d &= 2.375 - \left(\frac{INT}{ATT} \times 25 \right) && \text{YARDS = Passing yards} \\
 &&& \text{TD = Touchdown passes} \\
 &&& \text{INT = interceptions}
 \end{aligned}$$

$$\text{PasserRating}_{NFL} = \left[\frac{mm(a) + mm(b) + mm(c) + mm(d)}{6} \right] \times 100,$$

where $mm(x) = \max(0, \min(x, 2.375))$

All inputs to the formula are rather unequivocal numbers (e.g.: ATT, number of passing attempts), which make the metric sound and suitable for the purposes of comparing different players. However, there are some obvious limitations. PR rates passers rather than quarterbacks: all input variables are related solely to passing statistics of a player in a given game. The metric does not reflect some of the other important, unquantifiable qualities of a quarterback, such as leadership, play calling, timeliness, and performance during crucial situations. Similarly, it does not incorporate running ability, which has become an increasingly more important quality for quarterbacks over the recent years. Hence, although mathematically and statistically valid, the metric cannot be considered comprehensive when assessing the overall value of a quarterback.

Quarterback Rating: In 2011, ESPN released QBR⁸, a new metric intended to be a more meaningful alternative to PR, due to its ability to consider a variety of different aspects – i.e., not only passing ability – that contribute to the total rating of a quarterback. Unfortunately, the details of the formula have not been made publicly available to this time. According to ESPN, QBR was developed to measure the degree to which a quarterback contributed to scoring points for the team, and also the degree to which he contributed to the team's final victory. In order to achieve this, the metric incorporates the contexts and details of all

throws and what they mean for wins. A successful first down pass in a QB's own half with 10 seconds left to go is unlikely to lead to any points for his team; however, if the team is leading, such a pass increases the probability of winning (by letting the clock run out). In order to weigh different passes differently, a *win probability* function was developed by ESPN (requiring the analysis of each play in any NFL game over the previous ten years). This function is used to compute the "clutch index" for each play made by the quarterback, which is a measure of how great of an impact on winning or losing a game the play had. QBR also introduced the use of *expected points*: for each single play, there is an associated expected point gain depending on down and distance to touchdown. Furthermore, in QBR, credit is divided among teammates on the basis of how well they performed in a given play. In the end, the QBR is obtained by taking the weighted average of the points gained percentage for each play (out of its total expected points). The weights for the different plays are their clutch indices. The value of QBR is then normalized from 0 to 100.

3. History of Big Ben - Ben Roethlisberger

Ben Roethlisberger was drafted in 2004 by the owners of the Pittsburgh Steelers as one of the top prospects at the quarterback position. Led by his solid play, his team would go on to win the Super Bowl in only two years. However, he was involved in a motorcycle incident during the off-season, and he required surgery for facial fractures since he was not wearing a helmet. Although it was not illegal under Pennsylvania law, the decision to forgo a helmet drew fire from those who felt a quarterback should avoid dangerous activities. Contracts often include clauses preventing athletes from risky behavior that could jeopardize their playing ability⁹. Roethlisberger would go on to win a second Super Bowl in 2009 and established himself as one of the top quarterbacks in the league¹⁰. In March 2010, Roethlisberger was alleged of sexual assault in two separate instances. In Milledgeville, Georgia, a 20-year-old woman accused him of sexual assault at a nightclub. He allegedly told his bodyguards to guide her into a bathroom with him¹¹. The bodyguards then kept her friends from checking on her and prevented her from leaving before they had forceful intercourse. He was also accused of groping and penetrating a woman without consent in Harrah's Lake Tahoe two years earlier. Because of his friendship with the hotel's president, he threatened to have a female employee fired if she did not cooperate or stay silent afterward¹².

4. The Big Ben Case: Statistics, Misconduct and Decision Makers

The use of statistical information and numerical models in sports decision-making can be quite useful. A quantitative approach has replaced the "gut feelings" in sports like baseball with regards to signing and drafting players. Numeric values can be helpful in evaluating performance, but in the Roethlisberger case, decision-makers encounter problems that have nothing to do with on-the-field play. Roethlisberger's off-the-field conduct landed him in a considerable amount of legal trouble as well as NFL suspensions for conduct unbecoming of a professional football player.

Leading to Roethlisberger's 2010 sexual assault accusation that almost ended up in his release from the Steelers, organization was surrounded by information not accessible to all parties involved and had little to do with Roethlisberger's statistics. The decision-makers were in the Steelers front office: Owner, President, and General Manager. These decision-makers were influenced by outside pressures like the NFL, fans, coaches, players, and the press. The outside pressures were partial stakeholders as well, but those who most invested in the decision to sign or release Roethlisberger were the Steelers organization and the fans. Ultimately, the Steelers organization is the only decision maker. Due to Roethlisberger's off-the-field conduct and not his performance, which had been quite consistently good, the Steelers front office was forced to make a tough decision. They had signed the young Roethlisberger to a \$22.26 million deal in 2004 for six years. After winning Super Bowl XLII in 2008 he was given a new contract for eight years and \$102 million. In fact, after signing, he declared that he wished to be a Steeler for life.

Leading up to his misconduct in 2010, his statistics were among the best in the NFL. Before the 2010 season, however, the NFL commissioner suspended him for four games owing to his off-season misbehavior. The Steelers front office faced a tough decision. After signing Roethlisberger to an eight-year \$102 million contract they would legally owe him almost \$77 million if they released him. There were no legalities preventing Roethlisberger from collecting his contract money, so the Steelers front office had to decide if his misconduct warranted a release along with that sum of money. The Steelers organization prides themselves on being one of the NFL's classiest and most prestigious teams in the league's history. The Rooney family

owned the Steelers from the outset of the NFL and held player conduct higher in regards than on-the-field performance. The decision-makers had some raw numbers to make a decision and below are Roethlisberger's relevant statistics prior to his signing¹³.

SEASON	RANK (QBR)	TOTAL QBR	PR	RECORD	PLAYOFFS
2008	24	46.8	80.1	12 – 4	Won Super Bowl XLII
2009	8	67.3	100.5	9 – 7	Did not qualify

Roethlisberger's statistics were decent, but the second Super Bowl victory (his first was Super Bowl XL in 2005) was pivotal in setting him up for the 2008 deal. The Steelers saw potential in the young quarterback and hoped he could lead them to more championships. However, the 2010 sexual assault cases presented major obstacles. The NFL, fans, coaches, team, and press constantly pressured the front office to make a response to these cases. The Steelers also owed Roethlisberger a large sum of money that could handicap their signing capabilities with respect to salary caps by releasing him. The NFL voiced its opinion by suspending him for four games as they deemed his conduct unacceptable. Steelers fans were also torn, as they enjoyed watching Roethlisberger lead their beloved Steelers to their sixth Super Bowl but they did not feel his conduct embodied the Steeler organization. Opposing fans enjoyed making shirts and signs deriding him. The coaches and teammates on the Steelers were in consensus, besides a few dissenters, that Roethlisberger was their leader and could be remediated. The press enjoyed writing about the scandal and daily drama surrounding the Steelers' quarterback, which drew more attention to the situation. Roethlisberger became a dinnertime conversation in Pittsburgh and throughout the US via sports shows like Sports Center and Fox Sports. Such unquantifiable factors may appear irrelevant, but they have long-standing ramifications for decision-making for the Steelers.

The Steelers front office then decided to gather information that only they were privy to and engaged in many closed-door conversations with Roethlisberger. After a week of discussions, the owners, president, and manager were convinced Roethlisberger would repent for his actions. In his following season, he led the Steelers to another Super Bowl and was ranked in the top third of quarterbacks in his QBR. His QBR ratings remained consistently high in the following seasons as shown below.

SEASON	RANK (QBR)	TOTAL QBR	PR	RECORD	PLAYOFFS
2010	9	64.1	97.0	12 – 4	Lost Super Bowl XLV
2011	9	63.6	90.1	12 – 4	Lost Wild Card Round
2012	12	62.8	97.0	8 – 8	Did not qualify

What is most critical is the decision-making process and limited roles statistics can play in a scenario like the one with Roethlisberger. The Steelers quarterback performed well with respect to other NFL QBs in QBR, but his Super Bowl victory was an invaluable measure of success. A championship in professional football almost always guarantees a quarterback a hefty new contract. Deciding whether or not to release Roethlisberger had nothing to do with his performance or stats, but instead his off-the-field conduct. Money played a large role in the front office's decision-making process as it would handicap their organization, but they also had to bend to external pressures like fans, the NFL, and the press. The front office eventually benefited from information others did not receive, as they were able to incorporate the human factor into the decision-making process and convince Roethlisberger to reform his ways. As such, the story of Big Ben is just one of the many cases in sports where statistics alone is not sufficient for decision-making in the face of unexpected and extenuating circumstances coming from outside the sport.

IV. Conclusion

Professional sports are becoming increasingly intertwined with statistics and numerical models as different levels of a sporting franchise leverage relevant data and trends for better outcomes in decision-making. Similar to other real-world scenarios, decision-making in sports is dynamic and complex, and often subjected to influences that are unforeseen and injurious. The case of Ben Roethlisberger illustrates how a single major decision not only hinges upon statistics, but also on non-sports related factors like money, reputation and fanbases. As such, statistics in sports may further evolve to become more rigorous and accurate, but they will always be aggregated with other factors to determine the eventual outcome of decisions made.

Discussion Questions

1. To what extent are current numerical models in sports credible in their weighting and aggregation of metrics to assess the performance of an individual player or a team?
2. Should there be any enforcement or regulation in the manner of statistical analysis conducted in sports since some are imprecise or misleading?
3. Betting companies are heavily reliant on sports statistics and sports fans are often at the receiving end of uncertainties propagated by these numerical models. How can they best keep themselves informed?
4. It is often difficult to weigh the value of statistics among factors in decision-making for sports. How would different groups of stakeholders value the same piece of statistical information given to them?
5. How is decision-making in sports different from other policy modeling and assessment case studies?

References

-
- ¹ Robert P. Schumaker, Osama K. Solieman, Hsinchun Chen. *Sports Data Mining*.
 - ² Carl Bialik. The Count: Misconceptions About That Fourth-Down Call. 2009. [Online] Retrieved from <http://blogs.wsj.com/dailyfix/2009/11/20/the-count-misconceptions-about-that-fourth-down-call/>
 - ³
 - ⁴ Rich Lederer. Abstracts from the Abstracts. [Online] Retrieved from http://baseballanalysts.com/archives/2004/12/abstracts_from_22.php
 - ⁵ APBRmetric Central. [Online] Retrieved from <http://www.sonicscentral.com/statsite.html>
 - ⁶ Robert P. Schumaker, Osama K. Solieman, Hsinchun Chen. *Sports Data Mining*.
 - ⁷ Wikipedia – Passer Rating. [Online] Retrieved from http://en.wikipedia.org/wiki/Passer_rating
 - ⁸ Wikipedia - QBR. [Online] Retrieved from http://en.wikipedia.org/wiki/Total_quarterback_rating
 - ⁹ Judy Battista. *Roethlisberger Suspended for 6 Games*. 2010. [Online] Retrieved from <http://www.nytimes.com/2010/04/22/sports/football/22roethlisberger.html>
 - ¹⁰ Michael Fuoco. *Multiple injuries, few answers for Roethlisberger*. 2006. [Online] Retrieved from <http://www.post-gazette.com/stories/sports/steelers/multiple-injuries-few-answers-for-roethlisberger-437922/>
 - ¹¹ Russell Goldman. *Second Accusation Emerges Against Ben Roethlisberger*. 2010. [Online] Retrieved from <http://abcnews.go.com/GMA/TheLaw/ben-roethlisberger-accusation-emerges/story?id=10394816&singlePage=true#.UWWLiJO869M>
 - ¹² Jack McCallum. *The Hangover: Roethlisberger*. 2010. [Online] Retrieved from <http://sportsillustrated.cnn.com/vault/article/magazine/MAG1169185/1/index.htm>
 - ¹³ NFL Total QBR. [Online] Retrieved from <http://espn.go.com/nfl/qbr>

MIT OpenCourseWare
<http://ocw.mit.edu>

ESD.864 / 12.844J Modeling and Assessment for Policy
Spring 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.