Moneyball of X

Henri Hussinki1 and Harri Laihonen2

1 LUT University, Business School, Lahti, Finland
2 University of Eastern Finland, Faculty of Social Sciences and Business Studies, Department of Health and Social Management, Kuopio, Finland

henri.hussinki@lut.fi
harri.laihonen@uef.fi

Abstract: “Moneyball” (Lewis, 2003) is one of the most influential non-fictional books of the 21st century. It is a story of how a financially constrained baseball team, the Oakland Athletics, managed to turn around its fortunes and become competitive by exploiting an inefficient labor market with an extensive use of data analytics. Moneyball has had a great influence on how professional sports teams conduct their operations these days, it has spread from professional sports to several other realms, and its hypothesis has been a subject of academic studies. In the areas of business and economics, Moneyball has become a profound slogan for a firm’s data-driven revolution and the resulting new ways of conducting business. However, Michael Lewis, the author of Moneyball, accordingly thinks that the core message of Moneyball has been misinterpreted or misused to some extent (Dubner, 2022). It seems that adopting such a context-specific framework to other realms is neither easy nor simple, as the business logics and extra-organizational contingencies are different. The objective of this study is to better understand how Moneyball has been applied and discussed within academic research literature. This is done by systematically reviewing the Moneyball related literature that has been published in academic journals from 2003 to 2022. For both academic and managerial audiences, this study points out that sustained competitive advantage does not sprout from knowledge resources that are commonly available in the market. This study also serves as a reminder that a contemporary business and management buzzwords, such as Moneyball, Big Data, and Artificial Intelligence, may become elusive and misleading and the related research evidence scattered, if the academic research community cannot find ways to agree upon the definitions of the key concepts.

Keywords: Moneyball, Data analytics, Data-driven culture, Performance management, Literature review

1. Introduction

Moneyball offers a data-driven perspective to firms that may achieve competitive advantage. Firms that follow Moneyball philosophy are typically capable of identifying better data sources than their rivals and/or applying superior methods and models to analyze the data that is commonly available. Moneyball involves identification of new data or new ways of approaching the existing data, development of new insights based on data and rigorous analytics, and ultimately data-driven decision making (Hakes and Sauer, 2006). Moneyball is similar to business analytics that is defined as the extensive use of data acquired from diverse internal and external sources, statistical and quantitative analysis techniques, explanatory and predictive analytics models, and fact-based management to drive decisions and actions to proper stakeholders (Davenport & Harris, 2017; Soltanpoor & Sellis, 2016). However, the difference is that in the original Moneyball case example, competitive advantage was based on commonly available data and the first mover advantage: the Oakland Athletics was able to outperform their rivals by being the first one to identify baseball labor market inefficiencies through baseball match data and back then novel data analytics approaches. The market inefficiencies were caused by the rival teams’ misjudgment of some employee attributes and performance statistics that were believed to translate as team performance. For example, batting average and stolen bases were widely seen as among the most important statistics to assess players, but the Oakland Athletics were able to gain competitive advantage by looking at other statistics, such as on-base percentage (OBP, see Lewis, 2003).

In broader business management context, Moneyball can be defined as an extensive use of data-analytics to capitalize new business opportunities and to invent new ways of conducting business. Conversely, the increased use of data-analytics is not Moneyball if it does not lead to the capitalization of new business opportunities or new ways of conducting business. Neither is a new way of doing business Moneyball if it does not involve extensive use of data-analytics. Building on the knowledge-based view of the firm (KBV, see Grant, 1996; Spender, 1996), the benefits gained from the use of commonly available data in competitive markets are mainly short-term information benefits that are quite easily imitated by the rival firms. Commonly available data as such possesses almost zero intrinsic value, but when it is approached in an original way, analyzed rigorously, and its hidden insights are used in decision making, a firm may achieve competitive advantage. KBV posits that knowledge embedded in people is the firm’s most important strategic resource, and the firm’s task is to combine and coordinate knowledge to create value and sustained competitive advantage (Grant, 1996; Spender, 1996). In the original Moneyball case example the use of commonly available data was not the key to
Henri Hussinki and Harri Laihonen

the Oakland Athletics’ competitive advantage, but it was rather the people and knowledge on how to use the
data in a way that was never done before: 1) the new General Manager, a context expert, had the vision and
courage to make drastic changes to how employee performance was evaluated, 2) the data analytics expert
(hired by the new General Manager) had the suitable knowledge to view the baseball data in a new way that
produced information on undervalued employee performance statistics, and 3) the data-analytics knowledge
and methods that were adopted from other realms proved to work well in the professional sports context.
When these three different types of knowledge were combined, refined, and applied, it disrupted the Major
League Baseball’s labor market.

Moneyball hypothesis has been a subject of several academic studies, such as econometric evaluations (Hakes
and Sauer, 2006; 2007). However, according to the author of the original Moneyball book, Michael Lewis, most
of the Moneyball applications fail at their attempts to adopt the Moneyball philosophy to new realms, as they
try “to make some boring business much more entertaining to yourself than it actually is.” (Dubner, 2022) The
objective of this study is to review the academic Moneyball literature to understand whether and to what
extent the original idea of Moneyball has been correctly applied. By doing that, we show the merits of the
extensive use of data-analytics in transforming how business and operations are conducted and discuss why
firms need also other knowledge-related resources than data to gain sustained competitive advantage.

2. Methodology

Literature for this study were searched on January 30th, 2023, from citation and abstract databases Scopus and
Web of Science Core Collection. A single keyword “moneyball” was used and the search was targeted to
document titles, abstracts, and keywords, and limited to journal articles. The search in Scopus yielded 57
document results while the search in Web of Science returned 65 document results – altogether 122 results.
The citations and abstracts were then exported to a reference management service RefWorks. In RefWorks,
the 122 returned documents were checked for duplicates, which reduced the number of potentially relevant
documents to 88.

In the next step, the full texts of the 88 remaining documents were searched and downloaded. Full texts of 8
documents were not found, which reduced the number of potentially relevant documents to 80. Full texts of
the remaining documents revealed that 19 documents were not academic journal articles, which reduced the
document count to 61. Finally, out of 61 remaining documents, 24 were successful interpretations of the
Moneyball philosophy, while 37 were lacking some aspects of Moneyball. We will discuss both categories in
the results section.

3. Results

We analyzed 61 academic journal articles that were based, to a varying extent, on the Moneyball philosophy.
Our objective was to find out whether the various academic Moneyball studies were really about Moneyball or
something else. To determine whether a study was regarded as a successful interpretation of Moneyball, we
utilized the following qualification criteria:

• In the original baseball context: Moneyball is a combination of extensive use of data analytics and
exploitation of an inefficient labor market.

• In a broader management context: Moneyball refers to an extensive use of data-analytics to capitalize
new business opportunities and to invent new ways to conduct business.

By utilizing these definitions as qualification criteria, we identified 24 “Moneyball studies” and 37 “non-
Moneyball studies”. In the next sub-sections, we will describe our findings regarding these categories.

3.1 Moneyball Studies

Our analysis suggests that there were 24 such studies within the reviewed literature, which have successfully
adapted or adopted the Moneyball philosophy. These studies fall into three subsets regarding their contexts:
professional baseball (the original context), other sports, and non-sport realms.

3.1.1 Baseball context

Plethora of other Moneyball studies have been conducted in its original context, the MLB’s inefficient labor
market. In one of the most influential academic Moneyball studies, Hakes and Sauer (2006) focus on a case of
mispricing in the MLB labor market: the hitters’ (baseball player position) salaries in the change of the
millennium did not accurately reflect how different batting skills contributed to winning games, which allowed
one team, the Oakland Athletics, to gain sustained competitive advantage over their rivals. The competitive advantage evaporated many years later, as the other teams were able to close this knowledge gap. The same authors published another study a year later where they conducted a more refined analysis which suggested that this pricing anomaly existed well before the MLB’s “Moneyball period” and that the post-Moneyball market correction persisted (Hakes and Sauer, 2007).

Deli (2013) addresses the limitations of the earlier Moneyball studies and demonstrates how a methodological choice may improve understanding of the relative importance of the key predictors of baseball team performance. Brown et al. (2017) provides updated evidence that the post-Moneyball correction has persisted: the study demonstrates this persistence by showing that the OBP salary premium has risen for the class of baseball players known as “free agents” (i.e., players whose contracts have expired). With the expiration of their contracts, free agents are, in theory, involved in actively marketing their services to all MLB teams. Also, Congdon-Höhm and Lanning (2018) assess whether a broad range of production indicators (i.e., aggregate or rate statistics) translate into better salaries in the MLB’s free agency market and find that the Moneyball’s production indicators have accurately predicted the salaries in the 2000s. Conversely, Gin (2018) shows that the increasing reliance on the lower-paid baseball players has not translated as salary increase for the lower paid players. Interestingly, the analysis shows that the salary share for lower-paid players has even decreased in the Moneyball era. Holmes et al. (2018) builds upon Hakes and Sauer (2006; 2007), refines their earlier models, and finds new insights based on cleaner dataset: they find that the labor market for hitters/batters in baseball was efficient both before and after the appearance of Moneyball, thus providing new type of evidence. Pinheiro and Szymbanski (2022) arrive to the same conclusion about the efficient labor market for batters before and after the publication of Moneyball.

Moreover, Duquette et al. (2019) seeks to test the analytics underlying Moneyball by modelling MLB teams’ regular-season won-loss records over a long period (45 full seasons) and confirm the hypothesis of the significance of the previously overlooked predictors of team performance. They also find that some of these aspects are still undervalued by MLB teams in terms of the player-payrolls. With a slightly different focus, Caporale and Collier (2013) approach the Moneyball subject through the efficiency of the player draft in professional sports by evaluating a key implication of the Moneyball hypothesis, the superiority of college baseball players over high school players, which does not receive empirical backing. Finally, Abisaid and Cassidy (2017) investigate whether journalists have adopted sabermetrics, the use of advanced baseball statistics for making player projections and objectively measuring player performance, in their reporting of MLB stories since the print and theatrical release of Moneyball. They conclude that despite the increase of regular-season won-loss records over a long period (45 full seasons) and confirm the hypothesis of the lower- paid baseball players has not translated as salary increase for the lower paid players. Interestingly, the analysis shows that the salary share for lower-paid players has even decreased in the Moneyball era. Holmes et al. (2018) builds upon Hakes and Sauer (2006; 2007), refines their earlier models, and finds new insights based on cleaner dataset: they find that the labor market for hitters/batters in baseball was efficient both before and after the appearance of Moneyball, thus providing new type of evidence. Pinheiro and Szymbanski (2022) arrive to the same conclusion about the efficient labor market for batters before and after the publication of Moneyball.

3.1.2 Other sports

Moneyball has been adopted in other sports contexts, and association football is clearly the second most popular context for Moneyball studies. For instance, Gerrard (2007) evaluates the transferability of Moneyball philosophy to professional association football context and finds that there are technological, conceptual, and cultural barriers that make it rather difficult. Weimar and Wicker (2017) study whether certain association football players’ skills are undervalued, and they find that the labor market undervalues the player’s running distance. However, McHale and Relton (2018) find in their study that the amount of running is not positively related to team success. Gavião et al. (2020) develops a Moneyball model that is applied to evaluate the efficiency of association football player market, whereas Aydemir et al. (2021) provides a generic framework for estimating player performance and performing player-fit-to-criteria assessment, under different objectives, for left and right backs (association football player positions) from competitions worldwide. In addition, Zaytseva and Shaposhnikov (2022) conclude that there is a Moneyball effect in professional association football: defensive actions contribute positively to the team’s winning probability, but they are undervalued by the player market. Lastly, McHale and Holmes (2023) utilize advanced player rating systems to develop models for estimating the transfer fees of professional associate football players.

Stewart et al. (2007) apply Moneyball approach to study elite Australian Football. They use regression models to identify and quantify the relationship between important player statistics and match outcomes, suggesting that the Moneyball approach can be adopted to other team sport contexts, too. Millington and Millington (2015) argue that Moneyball is part of a highly important trend in sports: the utilization and impact of Big Data. Lastly, Plant and Stowe (2019) study horse racing and find an undervalued predictor of a racehorse’s late racing development and a measure that is valued by the auction market but predicts only early racing development (not career earnings).
3.1.3 Non-sport realms

Based on the reviewed literature, researchers have struggled to successfully adopt Moneyball to other realms than professional sports. There are, though, three instances where it seems to have worked. First, McHugh (2009) analyzes the applicability of the Moneyball book in management education. The study concludes that Moneyball is a useful complementary resource “to a traditional textbook when addressing topics such as organizational change and innovation, leadership, recruiting and selection, performance management, employee development, and negotiations” (McHugh, 2009, p. 219). Second, Mauboussin (2012) arguments for the applicability of Moneyball in various other realms but highlights that using the right statistics is prerequisite to compete and identifying and exploiting them before rivals do will be the key to seizing competitive advantage. Third, from the medicine industry’s point of view, Gayvert et al. (2016) proposes a model for clinical trials that goes beyond the drug-likeness methods that consider only the chemical properties of a molecule by including previously overlooked features related to the properties of the drug’s target.

3.2 Non-Moneyball Studies

37 of the reviewed articles did not qualify as successful applications of the Moneyball philosophy. These studies can be assigned to the following categories based on what they lacked: lack of data-analytics insight, and lack of new ways of conducting business.

3.2.1 Lack of data-analytics insight

As stated earlier in this study, Moneyball involves extensive use of data-analytics. However, it seems that a large proportion of the extant research that claims to be Moneyball has overlooked the role of extensive use of data-analytics. From the viewpoint of non-Moneyball studies in the sports context, Mason and Foster (2007) discuss the potential of applying Moneyball to a professional ice hockey league (National Hockey League). They highlight several issues that my hinder a widespread adoption of the Moneyball practices, but do not go as far as demonstrating how Moneyball would change the business operations of the National Hockey League teams. Moreover, Hughes et al. (2012) lists skill requirements for association football players, including seven different playing positions, but does not describe the data related to these "KPIs" and whether they statistically matter for team performance outcomes. Regan (2012) analyzes how different payroll efficiency strategies affect the baseball teams’ average game attendance but does not explain how extensive use of data analytics is linked to payroll efficiency. Elitzur (2018) analyzes the temporary and longer-term competitive advantage gained from establishing an analytics department within a baseball club, thus oversimplifying the Moneyball philosophy just as the existence of analytics department in a baseball club. Cassilo and Sanderson (2017) finds that media reports about analytics as the opposite of the conventional operational practices in the National Football League (American football) but does not explain how the National Football League teams could benefit from using data-analytics. Wang and Cotton (2018) examine how strategic and support teams’ experience is related to MLB team performance, and highlights the roles of social capital and human resource management. However, the study does not fully grasp the original idea of Moneyball. Kovalchik and Reid (2019) present a framework for sports-academia collaboration but does not present evidence of how it actually works. Roach (2022) study the impact of organizational decision-makers’ career concerns on their spending in the baseball player market and find that underperforming expectations decreases spending on signing bonuses. The role of an extensive use of data analytics is, however, very small.

In non-sports contexts, Wolfe et al. (2006) suggest that data has become a cornerstone of effective human capital management and therefore this should reflect also as changing competence requirements for human resource professionals. However, they do not show evidence of how extensive use of data-analytics improves the human resource operations. Wolfe et al. (2007) investigate the applicability of Moneyball in other realms outside baseball and conclude that the analytics-based performance assessment approaches might not be useful in those organizations with relatively low interdependence, but some of the Moneyball lessons, for instance, managing radical organizational change can be generalizable in broader context. In addition, Cullen (2009) takes 8 lessons learned from Moneyball to propose improvements for correctional agencies' operations. However, the study makes a comparison between Moneyball and evidence-based management, which are two quite different things. Antoun’s (2011) essay on the shifting forms of competition and changing political and cultural economies explains why poor teams and cities need to figure out something new to stay competitive. However, the connection to Moneyball philosophy is quite minimal. Evans (2011) suggests that Moneyball can and should be adopted in the healthcare sector to decrease the money spent on hospital readmissions, but the study does not provide any empirical backing for this idea. Kelly’s (2011; 2013) anthropology studies refer to Moneyball but only as a generic metaphor for change. Cohan (2012) suggests...
that science would benefit from utilization of advanced analytics methods, but there is no evidence of how these ideas work in practice.

Vito and Vito (2013) study the views of police managers on the applicability of Moneyball in policing. The findings suggest that statistical analyses are a promising lever to steer operations, to challenge the status quo, and to enable doing more with less. However, the elaboration on the use of data-analytics remains rather anecdotal. Butterworth (2014) tells a story about Nate Silver, who was able to predict the outcome of the 2012 US presidential election. Butterworth’s essay describes Nate Silver’s background, which involves some baseball analyst experience, but otherwise the study has nothing to do with Moneyball. Furthermore, Coe and Best (2014) proposes a model for the American Dental Association membership renewal but does not really show how the model improves the membership renewal in practice. Kaiser and Pratt (2016) uses data-analytics to show that pre-PhD publication activity predicts the researcher’s career publication activity. However, the novelty of this finding is questionable, as publication productivity is a commonly utilized criterion in academic recruiting.

Valerdi (2017) applies Moneyball to software project context and elaborates how it might help software teams to operate more efficiently and effectively through untapping their hidden value. However, the discussion remains on a rather hypothetical level. Norton et al. (2018) assess the extent of the marginal future payments for hospital payments but refers to Moneyball only in the introduction. Harewood (2019) tries to adopt the Moneyball philosophy to healthcare context, but the study focuses on basic data-analytics and established performance measures. Houck (2019) uses Moneyball metaphor to lay foundations to improve the performance of public forensic service providers but gives no empirical evidence. Howard and Mayes (2020) relate the employee credentials to their performance, with Moneyball mentioned only as a keyword. Szymanski (2020) connects dots between alchemy and sport analytics as they are both characterized by opacity and secrecy: evidence of success is limited according to the scientific criteria. Liu (2021) illustrates how social limits prevent mispriced human resources from being arbitraged away sooner, with implications for engaging cognitive diversity that go beyond sports. Mori et al. (2022) finds similarities between the sport and pharmaceutical industry: quantitative modelling could be useful for a broad set of stakeholders in the drug development technology ecosystem.

3.2.2 Lack of new ways of conducting business

Some of the non-Moneyball studies seemed to lack the insight of new ways of conducting business or operations. For instance, Tingling et al. (2011) study the National Hockey League player draft and find that earlier choice leads to better outcomes, but the effect becomes insignificant after the first 100 selections. However, this is common knowledge in the National Hockey League and based on talent scarcity. Bonomo et al. (2014) provides evidence that algorithms can be used to optimize management of a fantasy football team. Kringstad and Olsen (2016) study the effect of budgeted revenues to sporting outcomes, which is not a new way to analyze the input-output paradigm in professional sports. Kryscynski et al. (2018) study the individual level analytical abilities of HR professionals to their job performance, which is not related to Moneyball philosophy. Soland (2014) connect dots between Moneyball and predictive analytics in education — schools’ use of early warning systems to predict student dropouts. While this is an important research topic, its connection to Moneyball is rather vague. Ehrlich et al. (2021) mentions Moneyball but does not provide any definition or base its analysis on the Moneyball philosophy. Finally, Simmons (2022) recaps the influence of Moneyball on the rise of sports analytics on the input–output (pay-performance) relationship.

4. Discussion and Conclusion

The objective of this study was to review the academic Moneyball literature to understand whether and to what extent the original idea of Moneyball has been correctly applied. We find that Hakes and Sauer (2006, 2007) have had a great influence on the Moneyball studies conducted in the original baseball context. Some of the follow-up studies (e.g., Brown et al., 2017) confirm the inefficient labor market of the MLB and thus are aligned with Hakes and Sauer’s findings, while some other studies that utilize refined models and more (detailed) data show evidence of an efficient labor market (e.g., Holmes et al., 2018). Therefore, the research output is far from conclusive on the Moneyball’s mispricing hypothesis, which could be partially explained by Deli’s (2013) suggestion of the effect of the methodological choices on the research output. Those studies that point out to the post-Moneyball correction of the Major League Baseball labor market (e.g., Hakes and Sauer, 2007; Brown et al., 2017) receive backing from the KBV (e.g., Grant, 1996; Spender, 1996): when the key resources are commonly available (baseball match data) or easily transferred between firms.
(baseball data analyst leaves the organization and is employed by a rival), a firm cannot gain a sustained competitive advantage with those resources. In order to attain sustained competitive advantage, a firm should be dynamic in its data analytics by utilizing new and better data sources, combining new and existing data, renewing the data analytics skills and methods (e.g., machine learning), and developing new “research questions” to set objectives for data analytics.

Regarding the successful Moneyball adoptions in other sports contexts, it seems that association football (soccer) has been the most fertile soil for the Moneyball philosophy (e.g., McHale and Holmes, 2023). However, “flow team sports” such as association football are quite challenging contexts for Moneyball, as they are characterized with high interdependence between players, and it is difficult to isolate a single player’s effect on team performance. In non-sports contexts, we find that successful adoption of such a context-specific philosophy has been rare.

From the non-Moneyball findings, we can conclude that Moneyball as a concept has been used liberally. In some instances, it means just the existence of analytics department in organization (Elitzur, 2018), figuring out something new (Antoun, 2011), or just a metaphor to describe the increased opportunities related to data utilization (e.g., Vito and Vito, 2013). Some studies argue that Moneyball could be a good option to change how things are done in some contexts (e.g., Valerdi, 2017) but they do not provide evidence on how the extensive use of data-analytics provides benefits in practice.

All in all, those studies that have stayed true to the original idea of the Moneyball philosophy have been done in the baseball context or other sports context. At the other end of the spectrum, there are studies that have used Moneyball only as a metaphor without contributing to the Moneyball literature.

References


Henri Hussinki and Harri Laihonen


