The Quantitative Effect of Moneyball

WH299/399 Literature Review

Chris Hua

October 2, 2016

Introduction

This paper aims to determine what causal impact the publication of Michael Lewis's *Moneyball* (Lewis 2003) and the popularization of statistical techniques to determine player worth had on the business of sports.

In general, most people accept that there have been fundamental changes to how teams evaluate talent, as well as how they price players. There is no shortage of popular press giving anecdotal evidence that some 'Moneyball effect' exists (see e.g. Futterman (2011)), but relatively little work has been done to quantify this effect.

This paper intends to provide a broader review of the Moneyball effect, across leagues and across time, to determine the extent of the effect and when professional sports leagues began taking advantage of market inefficiencies.

Previous work

Moneyball (Lewis 2003) made the case that baseball teams only considered certain skills in their salary negotiations, and that these skills were not the ones most closely correlated with winning. The book itself told the story of Billy Beane, general manager of the Oakland Athletics, and The net effect is a market inefficiency in how players are signed and allocated. That is, baseball teams and their scouts overvalued particular 'big ticket' statistics, such as home runs or stolen bases, and ignored other factors with higher predictive power, such as slugging percentage (SLG) or on-base percentage (OBP).

This market efficiency was examined by Hakes and Sauer (2006) and found that a number of market adjustments were made in the MLB market, around 2003, when the book was published. Their core finding was that the labor market in MLB no longer exhibits the specific anomalies pointed out by Lewis. They compare salary distributions among a few particular populations that Lewis considered overvalued, such as hitters with over 25 home runs, and verify the impact of on-base percentage and slugging on win rates. They additionally emphasized the importance of knowledge diffusion in returning the market to efficiency, by noting the poaching of Billy Beane's assistants.

However, Hakes and Sauer (2006) examine only a very limited dataset: they focus on the years 2000-2004, on the Oakland A's, and on only a few player statistics. This is because their focus was narrowly set on confirming the ideas set forth in *Moneyball*. A broader approach might include a look at a more diverse set of player statistics, on all the teams, and on as many years as possible to evaluate a more accurate understanding of the Moneyball effect on the entire league. Their followup paper, Hakes and Sauer (2007), incorporates some of this criticism, but still includes very few years post-*Moneyball*, and still focuses on a smaller set of statistics.

Caporale and Collier (2013) also focuses on verifying one part of the Lewis *Moneyball* argument, that the A's focus on drafting college players was a more statistically sound process than drafting high school players. They do this by analyzing the relative impact that players had, depending on if they were drafted as high school or college players, and find no statistically significant support for college players outperforming high school players. More to our interest, they find no statistically significant change in draft positions for college vs high school players, meaning there was no effect from the publication of *Moneyball*.

There has been a fair bit of research around the application of Moneyball principles to other sports. Mason and Foster (2007) investigates this with respect to the National Hockey League. Hockey has had a very public, very contentious battle over the use of advanced analytics, particularly the Corsi number, a metric of shot attempts. In particular, the LA Kings and their manager, Dean Lombardi, won 2 championships in 3 years focusing on maximizing their Corsi score, which essentially meant picking players who dominate time of offensive possession (Gretz 2013). Similarly, the NBA has had a love-hate relationship with statistics, as many new coaches have experimented with advanced analytics, including the Philadelphia 76ers general manager, Sam Hinkie (Mase 2013), who was eventually ignomiously fired. The most promising use of analytics in basketball right now appear to be in making in-game decisions, such as teaching shooters to avoid long two-point jump shots, which offer exceedingly low risk-reward profiles (Ziller 2013). This shares some similarity with baseball and how Beane taught his players to be more patient and draw more walks to get on base. On a similar note, Michael Lewis profiled Shane Battier, and explained how Battier studied specific tendencies of Kobe Bryant and used those insights to shape his defensive positioning (Lewis 2009). In football, advanced stats have come to the limelight in the wake of fantasy football, including daily fantasy football. Fantasy football is a multi-billion dollar game, and has sprung up a cottage industry around predicting player performance (Wagstaff 2014). Additionally, (Weimar and Wicker 2014) applies Moneyball techniques to a European soccer league, by focusing on distance-run metrics of player performance to measure in-game effort and impact. (Wagstaff 2014) also notes that a computer science professor has focused on predicting soccer games with his team management algorithm, noting that soccer is easier for him to predict than other sports.

Possible extensions

There are a few ways in which my research can significantly contribute to the existing knowledge, deep and varied as it might be.

Broadly, we would expect that MLB teams get a better sense of which features to focus on when evaluating players over time. Lewis focused on on base percentage and slugging, and Hakes and Sauer (2006) shows that these underpriced attributes were quickly brought to an efficient position quickly. We would expect that teams have continued to search out other under-priced attributes of players, and that as time goes on, fewer underpriced features to remain untapped. The diminishing marginal returns of quantitative analysis to professional teams is quite possibly a concern, but more to the point of this paper, it would be valuable to explore which characteristics were discovered and when.

As a corrolary, it's likely that certain teams have 'better' (or luckier) quantitative analysts in their employ. In this case, we would expect different teams to take advantage of different pricing inefficiencies at different times; that is, a more quantitative team would exploit a pricing inefficiency before a less advanced team. By performing a temporal analysis, we could discover which teams have the most impactful analytics departments.

Additionally, we could examine what a measure of salary allocation efficiency might be. There are two avenues we should examine: how closely salaries paid line up with prior knowledge of player ability and output, and how closely salaries correlate with actual output. That is, comparing expectation and reality in salary vs performance. It would also be interesting to explore the success of trades, and what impact they might offer each party, in terms of discounted value.

Finally, a particularly rich topic of study in recent years is the application of Moneyball principles to other sports leagues. It would be interesting to apply a similar analytical framework to events in the National Hockey League or in the National Basketball Association. We discuss above the work of Mason and Foster (2007), but much work remains to be done with respect to hockey, basketball, and football. Of these, hockey and basketball are likely most fruitful because of the abundant data available, but for the same reason, researching the NFL may be a greater contribution to the literature because of the lower amount of research done already.

References

Caporale, Tony, and Trevor C. Collier. 2013. "Scouts versus Stats: the impact of <i>Moneyball</i> on the Major League Baseball draft." *Applied Economics* 45 (15). Routledge: 1983–90. doi:10.1080/00036846.2011.641933.

Futterman, Matthew. 2011. "Baseball After Moneyball." http://www.wsj.com/articles/SB10001424053111903791504576584691

Gretz, Adam. 2013. "How the Los Angeles Kings Corsi'd their way to the NHL's elite - CBSSports.com." http://www.cbssports.com/nhl/news/how-the-los-angeles-kings-corsid-their-way-to-the-nhls-elite/.

Hakes, Jahn K., and Raymond D. Sauer. 2006. "An Economic Evaluation of the Moneyball Hypothesis." *Journal of Economic Perspectives* 20 (3): 173–85. doi:10.1257/jep.20.3.173.

———. 2007. "The Moneyball Anomaly and Payroll Efficiency: A Further Investigation." *International Journal of Sport Finance* 2 (4): 177–89. http://search.proquest.com/docview/229341818?accountid=143123.

Lewis, Michael. 2003. *Moneyball: The Art of Winning an Unfair Game*. New York: W. W. Norton.

———. 2009. "The No-Stats All-Star." *New York Times Magazine*, February. http://www.nytimes.com/2009/02/15/magazine/ t.html?{_}r=1{\&}ref=magazine{\&}pagewanted=all.

Mase, Rick. 2013. "NBA embraces advanced analytics as Moneyball movement sweeps pro basketball." https://www.washingtonpost.com/sports/wizards/nba-embraces-advanced-analytics-as-moneyball-movement-sweeps-pro-basketball/2013/10/25/1bd40e24-3d7a-11e3-b0e7-716179a2c2c7{_}story.html.

Mason, DS, and WM Foster. 2007. "Putting Moneyball on ice?" *International Journal of Sport Finance* 2 (4). Fitness Information Technology: 206–13. http://ideas.repec.org/a/jsf/intjsf/v2y2007i4p206-213.html.

Wagstaff, Keith. 2014. "Fantasy Moneyball: How Algorithms Are Minting Big Winners." http://www.nbcnews.com/tech/innovation/fantasy-moneyball-how-algorithms-are-minting-big-winners-n193931.

Weimar, Daniel, and Pamela Wicker. 2014. "Moneyball Revisited: Effort and Team Performance in Professional Soccer." *Journal of Sports Economics*, December. SAGE Publications, 1527002514561789. doi:10.1177/1527002514561789.

Ziller, Tom. 2013. "Death by mid-range jumper: Who is starving themselves by feasting on long twos?" http://www.sbnation.com/2013/12/16/5215616/midrange-jumpers-shot-charts-wizards-bobcats-rockets.