Performance Persistence in Baseball Free Agency

WH299/399 Research Hypothesis

Chris Hua

24 October, 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.

We will focus on the question of 'performance persistence' in salary allocations.

Moneyball posited that certain statistics were undervalued by most baseball teams when considering contract negotiations, and that teams which properly valued those characteristics were able to reap outsized benefits. Previous research by Hakes and Sauer (2006) found that there was an inefficiency for the particular statistics that the Oakland A's used before 2004, and that after the publication of *Moneyball* and the increased transparency around those statistics, the market inefficiency was closed. Other teams adapted to what the Oakland A's used and increased how they valued those statistics.

However, even if these particular inefficiencies were played out, it is likely that other inefficiencies exist. 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.

Hypothesis

Major League Baseball teams can show performance persistence in acquiring undervalued players through free agency.

Target audience

This paper touches upon a number of interesting topics, and as such has a few potential audiences.

The core audience is researchers or practitioners interested in applying new research and frameworks to real world, applied problems. Obviously, some level of interest in sports and baseball would be beneficial, but at its core, the data analysis process is not limited to these fan bases.

The problem of performance persistence has most commonly been approached from a applied economics perspective, most typically in the finance literature. Brown and Goetzmann (1995) wrote the benchmark paper in performance persistence, finding that high performing mutual funds continue to do well, and poor performing mutual funds have higher likelihoods of disappearing. These findings and frameworks have been applied to a few other areas, including entrepeneurship (A. Gompers et al. 2008). The popularity of this framework in the finance context means researchers in finance may be interested in this research as well.

Notably, professional sports teams should be very interested in this analysis, for a few reasons. Beyond possibly better informing the team ofthe importance of data analysis, it is also possible that specific teams can be identified as being better at picking talent. These teams should be emulated, or if you can't beat

them, steal their statistical teams.

Significance

While performance persistence and "winners win" are well known in other contexts, little research has been done to apply these concepts to sports, and even fewer to sport contract analysis. This paper could be a significant contribution in these areas, and offer original analysis and research approaches to these questions of interest.

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*.

References

Brown, S, and William N. Goetzmann. 1995. "Performance Persistence." *The Journal of Finance* 50 (2): 679. doi:10.2307/2329424.

Caporale, Tony, and Trevor C. Collier. 2013. "Scouts versus Stats: the impact of Moneyball 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

Gompers, A, Anna Kovner, Josh Lerner, David S Scharfstein, Paul A Gompers Anna Kovner Josh Lerner David S Scharfstein, Paul Gompers, and David Scharfstein. 2008. "Performance Persistence in Entrepreneurship."

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.