International Journal of Sport Finance, 2018, **13,** 141-155, © 2018, West Virginia University

Moneyball and the Baseball Players' Labor Market

Paul M. Holmes,[1] Rob Simmons,[2] and David J. Berri[3]

[1]*Ashland University*
[2]*Lancaster University Management School*
[3]*Southern Utah University*

Paul M. Holmes is an assistant professor of economics at Ashland University. His research focuses on labor issues in sport, primarily on salary determination, dismissals, and discrimination.
Rob Simmons is a professor of economics at Lancaster University Management School. He has published widely on various topics in sports economics. He is a past president of North American Association of Sports Economists and is currently co-editor of *Economic Inquiry.*
David Berri is a professor of economics at Southern Utah University. His current research focuses on the economics of sports, specifically the topics of managerial decision-making, women in sport, and worker productivity.

Abstract

Michael Lewis's best-selling book, *Moneyball,* demonstrated the efforts of Oakland A's General Manager, Billy Beane, to create a successful baseball team in spite of its location in a small market. Previous studies have argued that the salary returns to the neglected skill of on base percentage (OBP) should rise once the Oakland A's hitters demonstrated proficiency with this skill. Our key result is that after *Moneyball* was published in 2003, hitter salaries for free agents signing new contracts were not more closely related to OBP. Consistent with efficiency, we find no long-term change in valuation in OBP. In contrast, we do find evidence of a rise in salary returns to productivity in the form of bases per hit ('power hitting') but this again is consistent with efficient market adjustment. In sum, it appears the labor market for hitters in baseball was efficient both before and after the appearance of *Moneyball*.

Keywords: efficient markets, moneyball, major league baseball

Introduction

Michael Lewis's 2003 book *Moneyball* (Lewis, 2003) describes the efforts of Oakland Athletics general manager Billy Beane to bring statistical methods to bear on the world of baseball management. By exploiting what Beane and his management group believed to be inefficiencies in player evaluation and baseball strategy, they endeavored to achieve output—wins—far in excess of what would generally be expected given their limited budget.

The central economic assertion in *Moneyball* is that baseball teams did not value players correctly. Beane believed that too much emphasis was placed on process—how players looked when they were playing—and not enough on the actual results. A devotee of baseball statistics guru Bill James, Beane attempted to implement the idea of results over process by seeking out players who did not fit the traditional description of a Major League baseball player, but could contribute to winning games via the fundamental aim of baseball: scoring runs, and preventing the opponents from scoring.

Most attention has been paid to the claim in *Moneyball* that hitters were being incorrectly evaluated. Specifically, it has been argued that on-base percentage (*OBP*, a measure of plate discipline, or the ability to avoid making an out when at the plate) was more important than slugging percentage (SLG, a measure of hitting prowess giving additional weight to hits that allow the hitter to advance "extra bases"). Though this claim wasn't new,[1] the book placed it unavoidably front-and-center.

The publication of *Moneyball* brought a slew of criticism from baseball insiders, much of it misguidedly at Billy Beane. But quietly, it appeared that baseball's general managers seemed to heed the ideas Beane had followed. In an article for *Sports Illustrated* in 2004, Lewis noted "The market for major league players with a high on-base percentage has tightened, thanks, in large part, to Oakland's success."[2] Formal econometric analysis (Hakes & Sauer, 2006) followed, indicating that indeed in 2004, the baseball labor market return to on-base percentage was estimated to be much higher than in previous seasons (i.e. before the publication of *Moneyball*); indeed, 2004 was the only year Hakes and Sauer could identify in which the return to on-base percentage was even statistically different from zero, at conventional levels of significance.

Sabermetricians and sports economists no doubt thought this would harken an era of enlightenment in Major League Baseball, with baseball labor contracts now correctly reflecting the relative importance of player skills and attributes. However Hakes and Sauer (2007) saw ominous signs: they estimated that the valuation of on-base percentage fell in 2005 and again in 2006, with its effect in 2006 no longer being statistically different from zero.

In this paper, we evaluate further what Hakes and Sauer dubbed "The *Moneyball* Hypothesis": that on-base percentage was undervalued prior to the publication of *Moneyball,* and implicitly that its publication—or at least, a contemporaneous market correction[3]—corrected this misvaluation. Using new-contract data, we are able to more accurately evaluate the labor market for hitters in baseball. The evidence we present suggests there has been no significant change to baseball's labor market. Contrary to the argument offered in *Moneyball,* this is because the market for hitters appears to have been efficient before and after publication of the book.

Related Literature

Moneyball's primary appeal to economists is in its implications about (a lack of) efficiency in pricing in the baseball players' labor market. As such a work, it was certainly

[1] Boston Red Sox General Manager Theo Epstein: "[The ideas Beane was using were] not a great secret, because Branch Rickey was using a lot of it a half century ago and Bill James had been writing about it for decades… [but the book] made it mainstream pretty quickly." In "The Art Of Winning An (even More) Unfair Game", *Sports Illustrated* September 26, 2011. Retrieved from http://si.com/vault/article/magazine/MAG1190632/index.htm on December 19, 2013.

[2] In "Out of Their Tree", *Sports Illustrated* March 1, 2004. Retrieved from http://si.com/vault/article/magazine/MAG1031308/1/index.htm on December 19, 2013.

[3] It has been argued that it wasn't necessarily the publication of *Moneyball* that triggered the change in the market, but the success of low-payroll teams like the Oakland A's.

not the first. Indeed, the article celebrated as founding the field of sports economics (Rottenberg, 1956) argued that the distribution of baseball talent was not affected by the *reserve clause,* and as such was efficient (i.e. the *invariance principle*). In the absence of free agency, rents would just flow to the teams owning the baseball players' contracts rather than the players themselves. With free agency, rents are shared between teams and veteran players in a bilateral bargain (Solow & Krautmann, 2010). Free agent players move to teams where their marginal revenue products and earnings are highest.[4]

Scully's pioneering and influential study (1974) attempted to estimate marginal revenue products of baseball players in the 1960s and 1970s. This was achieved, first, by regressing team revenues on wins produced by the teams and second, by regressing team wins on a set of player performance measures for hitters and pitchers. Scully computed estimates of rates of monopsony exploitation and found these to be substantial, especially for players subject to the reserve clause. Following on from Scully, Macdonald and Reynolds (1994) investigated baseball player salaries for 1986 and 1987 and concluded that, for free agents, 'Major League salaries generally coincided with estimated marginal revenue products." Hence, it appeared to these authors that experienced players were paid salaries roughly equal to marginal revenue product while young players were paid salaries below their marginal revenue products.

Krautmann (1999) took the Scully model a step further by assuming that free agents sought offers on *competitive* markets and so their salaries could be *assumed* to be equal to marginal revenue product. Using this approach, Krautmann estimated marginal revenue products for a sample of baseball players. Using the same approach, Krautmann, von Allmen, and Berri (2009) found substantial rates of monopsony exploitation for young players subject to reserve clauses in all four major North American sports leagues (NBA, MLB, NFL, NHL) in the 2000s. Of course, the assumption that free agent players are paid salaries according to competitive conditions can be questioned and needs to be tested, especially as bargaining and matching theories imply that even free agent salaries will lie below marginal revenue product. Bradbury (2013) offers several criticisms of Krautmann's revisions of the Scully model.[5]

Empirical research on baseball salary determination tends to follow the Mincer-style approach of regressing log salary on measures of age, experiences, player and team characteristics and measures of player performance (Bradbury, 2007; Brown & Jepsen, 2009; Hoaglin & Velleman, 1995; Holmes, 2011; Krautmann, Gustafson & Hadley, 2003; Link & Yosifov, 2012; Marburger, 1994). The most relevant articles to our paper are Hakes & Sauer (2006, 2007). In the first, the authors regress the

[4] However, it does not follow that the invariance proposition will actually hold in practice. The current system of baseball player free agency, where unrestricted free agency becomes an entitlement after six years of service, was initiated in the late 1970s. Schmidt (2011) argues that both the extent of player mobility and the characteristics of players who moved were affected immediately after the introduction of free agency, which appears to contradict the invariance proposition. Schmidt's results are consistent with Hylan, Lage, and Treglia (1997) who also argue that player mobility increased after free agency was introduced. Our analysis sits within the current free agency structure with no major changes in rules of entitlement.

[5] Salary data availability tends to restrict these papers to short panels. An exception that uses a longer-run perspective is Haupert and Murray (2012). These authors have salary and performance data for 400 players in a sample covering 60 years. These authors find that a closer relationship between pay and performance emerged after the Great Depression and before the onset of free agency in 1976.

logarithm of annual salary for MLB position players on a selection of explanatory variables, most relevant of which were OBP and SLG, for the years 2000–2004. In the second paper, the authors continue to explore this model, extending their data set to the period 1986–2006. Further support for the proposition that returns to OBP increased after publication of Moneyball has recently been offered by Brown et al. (2017) who also propose that the increased returns to OBP is confined to free agents rather than reserve-clause and arbitration eligible hitters.

Hakes and Sauer (2007) further explored the relative reward to various hitting skills by replacing OBP and SLG with "Bat" (traditional *batting average,* a measure of overall hitting prowess), "Power" (*isolated power,* defined as total bases per hit), and "Eye" (the ratio of walks and hit-by-pitch per plate appearance). Hakes and Sauer find that the coefficients in their salary equations associated with OBP/Eye increase in 2004 (in the first article) and in 2004–2006 (in the second article).

Model and Data

In this paper we seek to improve upon the modeling procedure used by Hakes & Sauer (2006, 2007), taking advantage of our modeling technique and a cleaner data set to investigate more accurately the *Moneyball* hypothesis. Hakes and Sauer used annual salaries for all position players (free-agents and not) in order to estimate their salary equations; we believe this is the wrong approach for this purpose and limited their ability to correctly identify the phenomenon they were examining. By including players on long-term contracts signed prior to a given year, and also players not yet eligible for free-agency, the link between salary and performance variables in the Hakes and Sauer data set is uncertain at best.[6] The salaries of free agents and reserve clause players are determined by separate processes (individual versus union negotiation) and with very different outside options. Hence, the salary returns to performance metrics are expected to be different for the two types of player not least because performances of reserve clause players have yet to be revealed. We believe that the intercept-shift approach adopted by Hakes & Sauer (2006, 2007)—allowing only for percentage differences in salary between otherwise-similar free-agents and restricted players—constitutes too strong an assumption about the salary-generating processes of the respective groups. This assumption becomes even worse when some players sign contract extensions, requiring forward-going salary to reflect not only expected future performance, but extant contract obligations. We focus on free agents, not only because this data is cleaner, but because this group in particular is generally thought to have their salaries determined by free-market forces; hence free agents ought to form the best test group for an examination of market efficiency.

The further problem arises for the Hakes and Sauer model that salary pre-determined by negotiation at $t - n$ is regressed on performance observed at a later period t. We believe this is unsound: previously-negotiated salary is based on future performance expectations. But current performance measures these expectations only loosely, adding substantial "errors-in-variables" statistical problems. Further, given

[6] Burger and Walters (2008) present evidence from the late 1990s to show that baseball teams systematically overvalued free agent players with more inconsistent performances, suggesting the presence of a winner's curse in bidding for free agents. A substantial part of Billy Beane's mission at Oakland A's was to improve the evaluation of veteran free agents in addition to improving the principles involved in trading players and in draft selection by using statistical analysis.

inconsistency of player performance due to form, luck, and injuries plus contributions of teammates and opponents, it is unreasonable to conjecture that performance observed at time of signing a contract is a precise predictor of future performance over the lifetime of a contract. Particularly, the link between salary and performance is likely to become more and more strained as the time between contract-signing and performance lengthens.

This problem is recognized by Brown et al. (2017) who regress free agent salaries on performance averaged over previous two seasons. In our setup, salary determined by negotiation at *t* is regressed on performance averaged over *t*-1, *t*-2 and *t*-3. We believe this approach to be more logically sound: current salary is determined by expectations about current and future performance formed using past performance, so modeling current salary with past performance seems inherently more reasonable.

Even worse, the modeling approach adopted by Hakes and Sauer (2006, 2007) artificially inflates their statistical significance levels, since they count the same player with the same multi-year contract as a separate and independent observation in multiple years. While using all players (free agents and not) is at least somewhat defensible when considering only a single season, it becomes substantially less so when comparing multiple seasons of data.

We elect to model only free agents; that is, only players who sign a new contract in a given year are included as observations for that year. By doing so, we feel confident that our model more accurately estimates the link between salary and performance.[7] In particular, we conjecture that the concentration on newly-signed contracts is critical in investigating any change in the influence of particular aspects of performance (such as on-base percentage) from year to year.

This leads to our choice of unit of observation: an observation in our data is a free-agent hitter (excluding pitchers, catchers, and designated hitters) who signs a contract from 1997–2012. We begin with a model superficially similar to Hakes and Sauer (2006): ln (*Salary*) is the dependent variable, and independent variables include *On-base percentage, Slugging percentage, At-bats, Age,* and *Speed*. Each of these are taken at their three-year averages, as we found greater explanatory power for salary using three-year averages than with one-, two-, or four-year averages. We include indicator variables for fielding position: *first base, second base, third base, short stop,* and *outfield* with *center field* as omitted category. From previous literature (Bradbury, 2007) we expect a positive salary premium for short stops and salary penalty for outfielders. For *Salary*, we use a player's average annual salary over the contract they just signed, adjusted for (CPI) inflation.[8] We added any signing bonus divided by the length of the contract. The other variables are described, along with summary statistics, in Table 1.[9]

[7] Hakes and Sauer were aware of this potential criticism of their approach, noting in Hakes and Sauer (2006) that they chose to "economize on data collection at the potential expense of precision" (first footnote, page 177).

[8] Hakes and Sauer (2006, 2007) instead used salary earned in the given year, and included a linear annual trend to account for inflation, implicitly assuming that inflation (or perhaps more correctly, expected inflation) was constant over their sample period.

[9] Data were obtained from the Sean Lahman database: http://www.seanlahman.com/baseball-archive/statistics/

Table 1. Independent Variables with Explanations and Summary Statistics

Variable	Description	Mean	Standard Deviation
OBP	On-base percentage from previous three seasons	0.337	0.029
SLG	Slugging percentage from previous three seasons	0.420	0.061
Avg	Batting average from previous three seasons	0.268	0.023
Eye	Walks plus hit-by-pitches divided by plate appearances from the previous three seasons	0.109	0.040
Iso	Isolated power (total bases per hit) from the previous three seasons	0.152	0.054
PA	Plate appearances in previous season	413.0	170.8
Age	Player's age at end of season	33.13	3.27
Speed	Bill James' Speed Score† from previous season	5.228	0.705
Position	Indicator variables: first base, second base, third base, shortstop, outfield	-	-

† Sabermetrician Bill James devised a series of box-score-based measures of player speed, based on stolen-base percentage, stolen base attempt rate, triples per hit, runs per time on base, and avoidance of grounding into double plays. Following James' lead, we average the top four of these in the previous year for each player, and call that the player's Speed Score. We normalize each individual speed component to an average of 5 and standard deviation of 1. More details here: http://www.baseballthinkfactory.org/btf/scholars/levitt/articles/speedscores.htm

The data yield an unbalanced panel: 16 years of observations (contracts signed after the seasons 1997–2012) with 32-70 observations per year (793 in total). 227 observations are of the only contract signed over this period by a given player; the maximum number of contracts signed in this period is seven (Rey Sanchez).[10] Our statistical model is standard OLS with year and team fixed effects. Standard errors are clustered by player since some players sign more than one contract. Results from our initial baseline model are presented as Model 1 in Table 2. All coefficients are highly significant at conventional levels and have the expected signs: players generally receive higher salaries when they are better hitters, younger, faster, and when they play more, *ceteris paribus.* Amongst the fielding positions, there are salary premia for third base and short stop, with the latter category especially valuable.

Did a Post-*Moneyball* Market Correction Take Place?

We are now in a position to investigate the "*Moneyball* hypothesis." We introduce an indicator variable for the post-*Moneyball* period (2004 onward) and interact it with the *OBP* variable. If Hakes and Sauer (2006) are correct in their assertion that the compensation for superior *OBP* in the market for baseball talent corrected after the publication of *Moneyball*, we expect the coefficient on this interaction term to be significant and positive. Results from the estimation of this model are shown as Model

[10] We do not adjust for players signing multiple contracts other than by clustering standard errors by player. Working with a subset of our data, Holmes (2011) explored the impact of players signing multiple contracts on regression coefficients in several ways, and concluded that an unadjusted OLS approach does not significantly alter coefficient estimates.

Table 2. Salary Regression Model Results with Post-Moneyball Interactions

	Model 1		**Model 2**		**Model 3**	
Variable	*Coefficient*	*p-value*	*Coefficient*	*p-value*	*Coefficient*	*p-value*
OBP	4.745***	0.000	5.903***	0.000	5.748***	0.000
SLG	6.315***	0.000	5.257***	0.000	4.742***	0.000
*OBP*Year≥2004*			-2.939	0.117	-2.704	0.165
*SLG*Year≥2004*			2.571***	0.010	3.748***	0.001
PA	0.0027***	0.000	0.0027***	0.000	0.0029***	0.000
Age	-0.021***	0.010	-0.020**	0.011	-0.020	0.103
Speed	0.154***	0.000	0.160***	0.000	0.071	0.224
Sample size	793		793		793	
R^2*(adj)*	0.695		0.696		0.705	

*, **, and *** indicate significance at 10%, 5%, and 1% levels. Models include position dummies and year and team effects.
Model 3 includes interactions between *Year≥2004* and *PA, Age,* and *Speed.*

2 in Table 2. The coefficient on the interaction term *OBP*Post 2003* is not significant. *Moneyball* had no lasting effect on the valuation of *OBP*. Hakes and Sauer did not consider interaction terms other than for *OBP.* Model 2 in Table 2 also includes an interaction term between *SLG* and the post-*Moneyball* period. This term has a statistically significant positive coefficient. Model 3 in Table 2 contains interaction terms for all variables and this delivers similar results to Model 2. However, of the control variables, only speed delivers an interaction effect that is significant at the conventional level (coefficient = 0.170, p = 0.024), and we prefer Model 2. We note that if we drop all control variable interaction effects apart from speed then the p value of the speed interaction term rises to 0.084. It appears that returns to slugging, as a standard measure of hitting prowess, actually increased after 2003. As far as we can tell, the discussion over *Moneyball* has not picked up this intriguing result. The insignificance of the *OPB post-Moneyball* effect remains whether we exclude team dummies or add player fixed effects.[11]

Why then did Hakes & Sauer (2006) find a post-*Moneyball* effect? To explore this, we introduce annual indicator variables and interact them with the *OBP coefficient.* One such model is shown in Table 3 as Model 4. In this model, we include indicator variables for each year 2002–2006. *OBP* has a statistically significant additional effect on salary *only in 2004*—the year after publication of *Moneyball,* and the same year Hakes & Sauer (2006) identified as reflecting a correction of the "mispricing in the baseball labor market" (*ibid.,* p. 174). This contrasts with the finding in Hakes & Sauer (2007) that *OBP* was significantly more highly rewarded in 2004 and 2005 than over

[11] Our results differ from Brown et al. (2017) in several respects. Their data set includes catchers. Their average performance measures are taken over two years rather than three in our models. They exclude team fixed effects. It is notable that they claim a significant increment to *OBP's* effect on free agent salary in the post-Moneyball period but find that *OBP* has an insignificant effect on free agent salary when considered on its own without the period interaction. However, we argue below that specifications based upon *OBP* and *SLG* should be criticised for collinearity between these two variables and below we propose what we regard as a superior specification based upon Hakes and Sauer (2007).

Table 3. Salary Regression Model Results with Year Interactions

	Model 4		Model 5	
Variable	*Coefficient*	*p-value*	*Coefficient*	*p-value*
OBP	4.711***	0.000	4.901***	0.000
SLG	6.285***	0.000	6.261***	0.000
*OBP*Year=2002*	2.667	0.312	8.253**	0.023
*OBP*Year=2003*	-2.116	0.450	-0.864	0.801
*OBP*Year=2004*	5.934**	0.014	1.918	0.471
*OBP*Year=2005*	-0.261	0.934	-2.457	0.501
*OBP*Year=2006*	-2.486	0.138	-5.537**	0.025
*SLG*Year=2002*			-4.148**	0.020
*SLG*Year=2003*			-1.421	0.538
*SLG*Year=2004*			2.921**	0.046
*SLG*Year=2005*			1.953	0.402
*SLG*Year=2006*			1.643	0.236
PA	0.0028***	0.000	0.0027***	0.000
Age	-0.020**	0.013	-0.020**	0.012
Speed	0.173***	0.000	0.144***	0.001
Sample size	793		793	
R^2(adj)	0.693		0.697	

*, **, and *** indicate significance at 10%, 5%, and 1% levels. Models include position dummies and year and team effects.

the rest of their time frame. Their 2006 estimate of the reward to *OBP* was also higher than average, though not significant. Hakes and Sauer were unsure how to interpret this: as "a hiccup, or if the market reflects a determination that teams were overpaying for plate discipline" (Hakes & Sauer, p. 186) in the years immediately prior.

We believe that the latter is correct. In the post-*Moneyball* period (2004–2012), *OBP* is not significantly more rewarded than pre-*Moneyball* (per our Models 2 and 3). Indeed, it doesn't show up as significantly more rewarded in any post-2004 year.

Why did Hakes and Sauer (2007) identify an effect still in 2005? We believe this is a consequence of their modeling approach of using all player-year observations. Many of the players in 2005 will have had their salaries determined in 2004 (by signing multi-year contracts that year); we conjecture that this dulled Hakes and Sauer's ability to pinpoint a one-period change in salary determination. Our free-agent data, however, can more sharply identify *when* changes occur; they show that the change (in the valuation of *OBP*) occurred in 2004, and was gone in 2005. Using Hakes and Sauer's term, it was 2004 that was the "hiccup."

There is a problem, though, with looking at the link between salary and these two measures of hitting performance. There is a strong correlation (0.557 in our data) between *OBP* and *SLG:* both are measures of hitting prowess. It is possible that it is a change in the valuation of *SLG* (as implied by our Model 2) rather than *OBP* that is driving the statistical significance result of *OBP*; omitting *SLG* may lead to omitted

Table 4. Correlation Matrix of Hitting Variables

	OBP	*SLG*	*Avg*	*Eye*	*Iso*
OBP	1	0.557	0.645	0.715	0.263
SLG		1	0.481	0.298	0.744
Avg			1	-0.044	0.126
Eye				1	0.356
Iso					1

variable bias, given the strong correlation between the two variables. To test for this, we interact our 2002 to 2006 indicator variables with both *OBP* and *SLG,* presenting the results as Model 5. Our model indicates a large and statistically significant increase in the valuation of *OBP* in 2002, with a corresponding decrease in the valuation of *SLG.* We suspect that this is a fluke rather than a real effect; taken together, and realizing that *SLG* generally takes higher values than *OBP,* the overall effect on salary in 2002 is relatively small. Contrast this with 2004: the returns to *OBP* and *SLG* both rise in 2004 (though only *SLG* is significant).

Our results so far suggest that changes in the reward for power hitting (measured by, for example, *SLG*) may have driven any post-2004 changes in the free agent hitters' market. To investigate this further, we pursue a similar approach to Hakes and Sauer (2007): splitting *OBP* and *SLG* up into three variables: *Avg, Eye,* and *Iso. Avg* is the three-year hitting average, intended to capture the common component of *OBP* and *SLG; Eye* is the three-year ratio of walks and hits to plate appearances, intended to capture the unique component of *OBP;* and *Iso* is the three-year isolated power (total bases per hit), intended to capture the unique component of *SLG.* The correlations of these five variables are given in Table 3; and as one can see, these factors are not highly correlated with one another.

In Model 6 in Table 5 we include these new variables in place of *OBP* and *SLG,* interacting them once more with the indicator variables for the post-*Moneyball* era. Model 7 includes the new variables interacted with the year indicators 2002 through 2006, In Model 6 we do not find a significant increment to the valuation of *Eye.* However, *Iso* does appear to receive an increment to its valuation in the post-2003 period (significant at 1 per cent). Salary elasticities with respect to *Iso* rise from 0.67 to 1.17 over the whole post-*Moneyball* period. In Model 7 we find a significant, positive coefficient for the *Eye*Year* interaction term for 2002 only, and this is followed later by a significant, negative coefficient for the 2006 interaction term. In contrast, we find a significant negative interaction term on *Iso*Year = 2002* followed by a significant, positive coefficient on *Iso*Year = 2004.* The lessons from Model 6 and Model 7 should be clear: one cannot confidently claim that it is the valuation of plate discipline rather than hitting power that drove any alteration of player valuation in the post-*Moneyball* period. Any increment in valuation of plate discipline is to be found at the beginning of Oakland A's success with low budget personnel but before the publication of *Moneyball.* This increment to salary returns to plate discipline was purely temporary. In 2006 we find a temporary *reduction* in salary returns to plate discipline.

The regressions in Table 5 do not include interaction terms for all years. To exploit the data over the whole sample period, we divide this into three eras: 1997–2002

Table 5. Salary Regression Model Results with Revised Performance Measures

	Model 6		Model 7	
Variable	*Coefficient*	*p-value*	*Coefficient*	*p-value*
Avg	13.883***	0.000	11.889***	0.000
Eye	2.866***	0.002	2.663***	0.000
Iso	4.530***	0.000	6.062***	0.000
*Avg*Year=2002*			5.733	0.132
*Avg*Year=2003*			-0.596	0.861
*Avg*Year=2004*			1.242	0.652
*Avg*Year=2005*			6.299	0.109
*Avg*Year=2006*			-2.312	0.455
*Eye*Year=2002*			5.349**	0.024
*Eye*Year=2003*			0.105	0.960
*Eye*Year=2004*			3.023	0.153
*Eye*Year=2005*			-4.547	0.170
*Eye*Year=2006*			-3.534***	0.007
*Iso*Year=2002*			-4.365***	0.007
*Iso*Year=2003*			-1.886	0.205
*Iso*Year=2004*			3.135**	0.030
*Iso*Year=2005*			1.510	0.500
*Iso*Year=2006*			1.899	0.183
*Avg*Year≥2004*	-2.851	0.130		
*Eye*Year≥2004*	-1.490	0.274		
*Iso*Year≥2004*	3.281***	0.001		
PA	0.0026***	0.000	0.0026***	0.000
Age	-0.021***	0.005	-0.021***	0.009
Speed	0.160***	0.000	0.141***	0.001
Sample size	793		793	
R^2(adj)	0.702		0.703	

*, **, and *** indicate significance at 10%, 5%, and 1% levels

before publication of *Moneyball,* 2003–2007 to show the immediate post-*Moneyball* period, and 2008–2012 for the recent period. We then create dummy variables for each of the three sub-periods and interact each of these with each of our three performance metrics. Regression results for this revised model specification are shown in Table 6 as Model 8. This specification smooths out the effects of individual years on pay-performance sensitivities. Model 8 reveals a *reduction* in marginal effect of *Eye* on free agent salary in the 2003 to 2007 period as compared to 1997 to 2003 and 2008 to 2012. In contrast, we find an increased sensitivity of free agent salary to *Iso* in 2003 to 2007 and 2008 to 2012 as compared to 1997 to 2003 (*p* values for F-tests of coefficient equality are 0.006 and 0.025 respectively, thus rejecting the null of coefficient equality). There is no significant variation in the effect of *Avg* on free agent salary across sub-periods. These results are reinforced by a specification with player fixed effects in which all *Eye* interaction terms are insignificant while *Avg* and *Iso* interaction terms remain significant, again with higher *Iso* effects for 2003 to 2007 and 2008 to 2012.

Table 6. Salary Regression Model Results with Period Interactions

	Model 8	
Variable	*Coefficient*	*p-value*
*Avg*1997-2002*	14.344***	0.000
*Avg*2003-2007*	11.126***	0.000
*Avg*2008-2012*	11.182***	0.000
*Eye*1997-2002*	2.891***	0.004
*Eye*2003-2007*	0.773	0.462
*Eye*2008-2012*	3.253**	0.015
*Iso*1997-2002*	4.597***	0.000
*Iso*2003-2007*	7.259***	0.000
*Iso*2008-2012*	7.119***	0.000
PA	0.0026***	0.000
Age	-0.021***	0.007
Speed	0.159***	0.000
Team effects	Yes	
Player effects	No	
Sample size	793	
R^2(adj)	0.700	

*, **, and *** indicate significance at 10%, 5%, and 1% levels

Did the Market Need to Change?

The discussion thus far has focused on how the market for hitters has changed. Perhaps a more important question is whether or not the market *needed* to change. To address this issue, we need to consider how the salary returns to the statistics tabulated for hitters relate to how those statistics relate to wins.

We begin with *OBP* and *SLG.* Table 7 reports the results of a regression of the log of runs scored per game by a baseball team on a team's *OBP* and *SLG.* The results—estimated on MLB team panel data[12] from 2000 to 2013—indicate that 92% of the variation in runs scored per game is explained by these two factors. Since this model is double-logged, the reported coefficients are the elasticity of runs scored to each factor. These results indicate that *OBP* has a bigger impact than *SLG* on runs scored per game.

When we look back at the aforementioned salary results, though, a different story is told. The elasticity of free agent salary on *OBP* is 1.95 while the elasticity found with respect to *SLG* is 2.68. So when it comes to salaries, *SLG* has the highest impact. This suggests the market for hitters in baseball is not perfectly efficient.

There is a problem, though, with this interpretation. As noted, *OBP* and *SLG* are highly correlated.[13] Consequently, we might suspect our results. To overcome this issue we turn to the trio of *Eye, Iso,* and *Batting Average.* The correlations between these

[12] These data can be found at Baseball-Reference.com.

[13] In this sample of team data, *OBP* and *SLG* have a 0.79 correlation.

Table 7. Team Runs Scored per Game Regression Model Results

	Model 9		Model 10	
Variables	*Coefficient*	*p-value*	*Coefficient*	*p-value*
OBP	1.213***	0.000		
SLG	0.980***	0.000		
Eye			0.254***	0.000
Iso			0.973***	0.000
Avg			1.841***	0.000
Constant	3.729***	0.000	4.143***	0.000
Sample Size	432		432	
R-squared	0.921		0.923	

Model is double-logged; robust standard errors
*, **, and *** indicate significance at 10%, 5%, and 1% levels

three factors are much lower.[14] When runs scored per game is regressed on these three variables we see that explanatory power is essentially the same (relative to what we saw with respect to Model 11). Furthermore, the estimated elasticities indicate that batting average has the largest impact on runs scored per game. This is followed by *Iso* and *Eye*. In sum, the ability to draw a walk has the smallest impact on salaries.

Turning back to our model of salaries we see essentially the same story. Using Model 6 the estimated elasticities of salary with respect to these three variables are as follows:

- *Batting Average*: 3.72
- *Iso:* 0.69 (1.20 in post-*Moneyball* era)
- *Eye:* 0.31

Looking at Tables 5 and 7, we see that these results conform in both order and essentially magnitude to what we see with respect to runs scored per game. Teams primarily pay for the ability to hit the ball. Power is also valued, while the ability to draw a walk is not as important. When we look at runs scored per game we see essentially the same story. In sum, the focus on a player's ability to draw a walk appears to be misplaced.[15]

Ironically, the attention that has been paid both in the media and in the literature to the singular *OBP* statistic was warned against in *Moneyball,* quoting baseball writer and statistician Bill James, who said that people had a tendency to "glom onto something superficial in the book and misunderstand its underlying message."[16]

[14] In this sample of team data, *Eye* and *Iso* have a 0.43 correlation; *Eye* and *Batting Average* have a 0.18 correlation, while *Iso* and *Batting Average* have a 0.08 correlation.

[15] Using a similar approach, Hakes and Sauer (2007) also concluded that the market reward for the three separable skills was largely in line with their respective contributions to winning, except they found that *Eye* became more highly rewarded post-2004, while we find that *Iso* is more highly rewarded instead. We suspect that this might have resulted from their inclusion of catchers in their data, a position that we excluded. Catchers typically have strong pitch-identification skills, allowing them to draw a lot of walks; we suspect that this somehow led their model to show different conclusions from ours.

[16] *Moneyball,* page 96. Quote from James (1988).

The central tenet of *Moneyball* was never that *OBP* was overarchingly important; it was that teams were better off looking at data rather than just trusting their eyes. In the early 2000s, Oakland General Manager Billy Beane believed players with high *OBP* were undervalued, but also that speed and fielding were overvalued, that college players in the draft were undervalued relative to high school players and that players who looked like baseball players "ought to look" were overvalued relative to unathletic-appearing players. Beane reversed several of these positions over the coming years, presumably as his interpretation of the data has changed.[17]

Conclusions

Michael Lewis's best-selling book, *Moneyball,* and the film based on the book demonstrated the efforts of Oakland A's General Manager, Billy Beane, to create a successful baseball team despite a small team payroll and a location in a small market in the San Francisco bay area. Hakes and Sauer (2006) argued that the salary returns to the neglected skill of on base percentage (*OBP*) should rise once the Oakland A's hitters demonstrated proficiency with this skill. The demand for ability to generate walks as a key component of *OBP* should rise and the salary return to this skill should permanently increase according to the authors as other teams would imitate the Oakland A's playing strategy and demand players who could achieve higher *OBP* values.

We offer radically different results and hence propose a different explanation of salary returns to hitting skills. Specifically, unlike Hakes and Sauer, we focus on the salary returns to observable skills of free agents since restricted players have salaries that are not determined on the open market. Our large data set consists of 793 player-seasons and spans a period before and after publication of *Moneyball*. Our key result is that it does not appear that the market for hitters in baseball has changed very much over time in terms of valuation of batting average or plate discipline. Furthermore, it also does not appear that any correction was necessary in this market in terms of valuation of these productivity measures. Despite the interpretation many people have offered with respect to the Moneyball story, the market for hitters in baseball appears to conform to what one would expect from an analysis of how runs are created in baseball.

The one change in the baseball hitter's labor market that withstands the scrutiny of our various estimations is a rise in the salary return to isolated power. This can be reconciled with efficient market adjustment. A relative scarcity of ability to convert hits into bases would drive up the price of isolated power. Data taken from baseball-reference.com show a trend reduction in team isolated power measure from 1997. Formally, a team fixed effects regression of team isolated power on time trend shows a significant, negative coefficient. Moreover, home runs per game fell over this time. From our free agent data set, average isolated power regressed on time trend also reveals a significant, negative coefficient. There are various reasons for increased scarcity of bases per hit including reduced use of performance-enhancing drugs by hitters as a result of a more rigorous testing and penalty regime introduced in 2005, better fielding and improved pitching performances which in turn might reflect improved

[17] For example, by 2010 Beane believed unpolished high-school players were a better bet than college draftees. See "The Revolutionary," *Sports Illustrated,* January 15, 2010; retrieved from http://sports.espn.go.com/espn/otl/columns/story?columnist=bryant_howard&id=4357166 on February 13, 2014.

pitching efficiency but also umpire calls that became more favorable to pitchers (Mills, 2014; Kim & King, 2014). A rise in salary return to isolated power might also reflect increased fan demand for home runs and in turn greater marginal *revenue* product from home runs. These factors take the analysis of the baseball hitters' labor market beyond consideration of plate discipline and are plausible reasons for variations in salary returns that are not connected to the use of sports analytics by statisticians hired by baseball teams and are compatible with efficient market adjustment to revealed information.

References

Bradbury, J. C. (2007). Does the Baseball Labor Market Properly Value Pitchers? *Journal of Sports Economics, 8,* 616–632.

Bradbury, J. C. (2013). What is Right with Scully Estimates of a Player's Marginal Revenue Product? *Journal of Sports Economics, 14,* 87–96.

Brown, D., Link, C. and Rubin, S. (2017). *Moneyball* after 10 years: How have Major League Baseball salaries adjusted? *Journal of Sports Economics, 18,* 771–786.

Brown, K., and Jepsen, L. (2009). The impact of team revenues on MLB salaries. *Journal of Sports Economics, 10*(2), *192-203.*

Burger, J., and Walters, S. (2008). The Existence and Persistence of a Winner's Curse: New Evidence from the (Baseball) Field. *Southern Economic Journal, 75,* 232–245.

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

Hakes, J. K., & Sauer, R. D. (2007). The Moneyball Anomaly and Payroll Inefficiency. *International Journal of Sport Finance, 2*(4), 177–189.

Haupert, M., and Murray, J. (2012). Regime Switching and Wages in Major League Baseball under the Reserve Clause. *Cliometrica, 6,* 143–162.

Hoaglin, D., and Velleman, P. (1995). A Critical Look at some Analyses of MLB Salaries. *The American Statistician, 49,* 277–285.

Holmes, P. M. (2011). New Evidence of Salary Discrimination in Major League Baseball. *Labour Economics, 18*(3), 320–331.

Hylan, T., Lage, M., and Treglia, M. (1997). Institutional Change and Invariance of Behaviour in Major League Baseball. *Applied Economics Letters, 4,* 311–314.

James, B. (1988). *The Bill James Historical Baseball Abstract.* Villard Books.

Kim, B., & Kang, M. (2014). Seeing Stars: Matthew Effects and Status Bias in Major League Baseball Umpiring. *Management Science, 60,* 2619–2644.

Krautmann, A. (1999). What's Wrong with Scully-estimates of a Player's Marginal Revenue Product? *Economic Inquiry, 37,* 369–381.

Krautmann, A., Hadley, L, and Gustafson, E. (2003). A Note on Structural Stability of Salary Equations: Major League Baseball Pitchers. *Journal of Sports Economics, 4,* 56–63.

Krautmann, A., von Allmen, P., and Berri, D. (2009). The Underpayment of Restricted Players in North American Sports Leagues. *International Journal of Sport Finance, 4,* 161–175.

Lewis, M. (2003). *Moneyball: The Art of Winning an Unfair Game.* Norton.

Link, C., & Yosifov, C. (2012). Contract Length and Salaries Compensating Wage Differentials in Major League Baseball. *Journal of Sports Economics, 13,* 3–19.

Macdonald, D., & Reynolds, M. (1994). Are Baseball Players Paid Their Marginal Products? *Managerial and Decision Economics, 15,* 443–457.

Marburger, D. (1994). Bargaining Power and the Structure of Salaries in Major League Baseball. *Managerial and Decision Economics, 15,* 433–441.

Mills, B. (2014). Social Pressure at the Plate: Inequality Aversion, Status, and Mere Exposure. *Managerial and Decision Economics, 35,* 387–403.

Rottenberg, S. (1956). The Baseball Players' Labor Market. *Journal of Political Economy, 64*(3), 242–258.

Schmidt, M. (2011). Institutional Change and Factor Movement in Major League Baseball: An Examination of the Coase Theorem and the Invariance Principle. *Review of Industrial Organization, 39,* 187–205.

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

Solow, J., & Krautmann, A. (2010). A Nash Bargaining Model of the Salaries of Elite Free Agents. *Journal of Sports Economics, 12*(3), 309–316.