Masthead Logo

Smith ScholarWorks

Mathematics and Statistics: Faculty Publications

Mathematics and Statistics

9-2014

Quantifying Market Inefficiencies in the Baseball Players' Market

Benjamin Baumer
Smith College, bbaumer@smith.edu

Andrew Zimbalist
Smith College

Follow this and additional works at: https://scholarworks.smith.edu/mth_facpubs

Part of the Statistics and Probability Commons

Recommended Citation

Baumer, Benjamin and Zimbalist, Andrew, "Quantifying Market Inefficiencies in the Baseball Players' Market" (2014). Mathematics and Statistics: Faculty Publications, Smith College, Northampton, MA.
https://scholarworks.smith.edu/mth_facpubs/24

This Article has been accepted for inclusion in Mathematics and Statistics: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu

Submitted to the Eastern Economic Journal

QUANTIFYING MARKET INEFFICIENCIES IN THE BASEBALL PLAYERS' MARKET

By Ben Baumer*

Smith College
bbaumer@smith.edu
and

By Andrew Zimbalist*
Smith College
azimbali@smith.edu

JEL codes: L83 - Sports; Gambling; Recreation; Tourism, J49 - Other

Among the central arguments of the bestselling book and movie *Moneyball* was the allegation that the labor market for baseball players was inefficient in 2002. At that time, Billy Beane and the Oakland Athletics used observations made by statistical analysts to exploit this market inefficiency, and acquire productive players on the cheap. Econometric analysis published in 2006 and 2007 confirmed the presence of an inefficient market for baseball players, but left open the question of to what extent, and how quickly, a market correction would occur. We find that this market had in fact already corrected by 2006, and moreover argue that the perceived market response to *Moneyball* in 2004 is properly viewed as part of a more gradual longer-term trend. Additionally, we use official payroll data from Major League Baseball to refute a previous observation that the relationship between team payroll and performance has tightened since the publication of *Moneyball*.

*This paper is adapted and updated from an analysis in our forthcoming book, *The Sabermetric Revolution: Assessing the Growth of Analytics in Baseball*, to be published in December 2013 by the University of Pennsylvania Press. (Baumer and Zimbalist, 2014)

Keywords and phrases: baseball, sabermetrics, market inefficiencies, labor markets, statistical modeling

1. Introduction. Michael Lewis's bestselling book *Moneyball* (Lewis, 2003) was published in June of 2003. Stripped of its storytelling, Lewis presented two strong theses. The first was that baseball executives had been using the wrong metrics to value the productivity of players and that this tendency was reversed by general manager Billy Beane and the Oakland Athletics in the early 2000s. The second, as a consequence, was that certain skills were undervalued (market inefficiency) and intelligent executives at small market teams could overcome their competitive disadvantage by exploiting the skill undervaluation.

While the Athletics performed well above what their meager resources would lead one to predict, the notion that sabermetric smarts could undo baseball's competitive balance problems engendered significant skepticism. If Lewis were correct in his assertion of skill undervaluation, then the world would know the secret shortly after his book was published and Billy Beane's advantage would soon disappear.

Central to Lewis's narrative was the importance of a player's walk rate. On-base percentage (OBP)[1] was more closely correlated with a team's win percentage and revealed higher consistency from one year to the next for individual players than did a player's simple batting average (BA). OBP was more closely correlated to win percentage because a walk not only put a runner on base and sometimes moved other runners up, but it also allowed an additional batter to reach the plate during an inning and wore down the arm of the opposing pitcher. Batting average did not capture the important skill of having a good batter's eye and being able to work a walk.

While the superior value of the OBP metric, relative to BA, is manifest upon a moment's reflection, sabermetricians have turned their attention to more ambitious metrics, such as quantifying a fielder's range or separating out the value of good defense from good pitching. The analytic work done in baseball front offices these days is guarded closely as proprietary with teams seeking to reap the benefits of their discoveries as long as possible before other teams catch on.

As the work of sabermetricians becomes more secretive, the speed of the market adjustment process may slow. In the case of OBP, however, once the Athletics' strategy was recognized, it was easy for other teams to emulate it. Indeed, it is possible that the appreciation of the skill of walking became overdeveloped and OBP became overvalued; that is, the return to the ability to work a walk may have outpaced the value of working a walk. Ironically, to the extent that the market overadjusts to OBP or to other skills, gen-

[1] OBP is equal to (*hits* + *walks* + *hit by pitch*)/(*at-bats* + *walks* + *hit by pitch* + *sacrifice flies*).

eral managers who pursue this skill in the marketplace will find that their team can be disadvantaged by the application of sabermetric knowledge. In contrast, the laggard general manager who eschews analytics may be temporarily advantaged by his obscurantism.

In any event, now, with ten years of market response since the publication of *Moneyball*, it is interesting to follow how the skill of plate discipline has changed in value over time. To study this question, we began with two papers by Jahn Hakes and Raymond Sauer (Hakes and Sauer, 2006, 2007). In a 2006 article, Hakes and Sauer employ data from 2000-2004 to compare the relative salary returns to changes in slugging percentage (SLG)[2] and on-base percentage (OBP) to the relative impact of SLG and OBP on win percentage. The key distinction they make is that SLG is a traditional measure of batting prowess, while OBP is the sabermetric variable of choice. They find that OBP is undervalued relative to SLG between 2000 and 2003, but that this undervaluation is abruptly reversed in 2004, the year after *Moneyball* is published. While instructive, this first article is limited by (a) only including five years of data and one year of data after the book's publication, (b) the fact that both OBP and SLG include singles in the numerator and outs in the denominator and, hence, are correlated with each other, and (c) the likelihood that the valuation of SLG is a function both of its contribution to winning and, independently, its contribution to fan enjoyment of the games (extra base hits and home runs are more exciting to watch than walks and singles.)

Hakes and Sauer published a second paper in 2007 that extended their data set to 1986-2006 and introduced a refined separation of different hitting skills. Hitting skills were now delineated as *Bat* (*batting average*), *Eye* ((*walks* + *hit by pitch*)/*plate appearances*) and *Power* (*bases per hit*). The basic results corroborated those of the 2006 study. *Eye* was relatively undervalued until 2004, when its valuation spiked. The authors also found that the returns to *Eye* diminished in 2005 and 2006, indicating a possible overcorrection in 2004. By 2006, the return to *Eye* or plate discipline was at the same level it had been in 2003.

In this paper, we seek to further refine the modeling of these relationships and to take advantage of the time passed to follow the pattern of market response through 2012.

2. Data & Methodology.

Like Hakes & Sauer, our primary data source is the Lahman database (Lahman, 2013), which contains seasonal statistics for every player in Major League Baseball history going back to

[2]SLG is equal to $(1 \cdot \textit{singles} + 2 \cdot \textit{doubles} + 3 \cdot \textit{triples} + 4 \cdot \textit{home runs})/\textit{at-bats}$.

1876. We used the 2012 version of the database, and focused on the 798 team-seasons that occurred between 1985 and 2012.

The Lahman database also contains information about player salaries in each year. However, the structure that governs the salaries of Major League players is notoriously complex. There are three major categories into which each player falls at the end of each season: free agent, arbitration eligible, and pre-arbitration eligible (a.k.a., 0-3 player)[3]. If a player is a free agent, then his salary in the next season is likely to closely reflect the market price for his services. Conversely, if he is pre-arbitration eligible, then his salary is unilaterally determined by his club, and is likely to be very close to the league minimum ($490,000 in 2013). If he is arbitration eligible, then his salary will be determined through a two-party arbitration process, which will likely yield something approximating a market rate. (For free agent and arbitration-eligible players, the correspondence between output and salary will also be distorted by the presence of long-term contracts.) Thus, knowing into which of these three buckets each player falls is an important factor in estimating his salary. The central piece of data that determines into which bucket each player falls is his major league service time, which accrues daily for each day the player spends on a major league roster. These data can be hard to find, but our service time data, which does not come from the Lahman database, allows us to make more precise determinations on this subject than Hakes and Sauer.

What follows is an overview of our methodology:

1. We build a model for labor market *productivity*, using three key offensive performance variables. This gives us an understanding of how these skills translate into winning on the field.
2. We build a corresponding model for labor market *valuation*, using the same three key variables. This gives us an understanding of how these skills are compensated on the labor market.
3. We compare the values of the corresponding coefficients. This helps us to assess inefficiencies in the labor market for baseball players.

3. Modeling Labor Market Productivity.

First, we want to understand how certain skills translate into winning. Specifically, we are interested in the following three skills:

1. *Eye*: walks plus hit-by-pitches per plate appearance
2. *Bat*: hits per at-bat (a.k.a., batting average)

[3] To be precise, under the most recent collective bargaining agreement, the 22 percent of players between two and three years in the majors with the most service time also qualify for salary arbitration. In previous agreements, 17 percent of such players qualified.

3. *Power*: total bases per hit (a.k.a., Slugging percentage divided by batting average)

These three skills are largely uncorrelated with each other, or at least, far less correlated than more comprehensive statistics like on-base percentage (OBP) and slugging percentage (SLG). While it may be the case that players with excellent plate discipline tend to hit for more power, there is no reason, a priori, to think that this might be the case. Moreover, unlike OBP and SLG, the pairwise relationships among these three variables contain no functional dependence.

¡¡INSERT TABLE 1 HERE¿¿

Our dependent variable is team winning percentage, but we consider this relative to .500, which defines an average team. Also, because baseball has a beautiful (though perhaps often overlooked) duality between offense and defense, every statistic that measures something good for the offense measures something bad for the defense. Thus, our explanatory variables measure the difference between each team's offensive performance and its defensive performance. Accordingly, we fit, using least squares, the regression model:

$$WPct - .500 = \beta_1(Eye - EyeA) + \beta_2(Bat - BatA) + \beta_3(Power - PowerA),$$

where $EyeA$, $BatA$ and $PowerA$ represent the opposing team's offensive performance, when playing the team in question. This is equivalent to the model fit by Hakes & Sauer, and our coefficients agree with theirs over similar time intervals. A summary of the results is shown in Table 1.

The first column of Table 1 shows the regression results over the full 28-year period in our study. We note that while there is clearly more to winning games than the three variables we have measured here, our model explains 81% of the variation in team winning percentage. In the four rightmost columns in Table 1, we have broken the 28 years into four distinct periods. The first (1985-1997) is far longer, and can be thought of as something of a control. The period from 1998-2002 contains the five years immediately preceding the release of *Moneyball* the book, along with the actual season about which the book was written (2002). The five-year period immediately following (2003-2007) is when we would expect to see the ideas from *Moneyball* implemented in the baseball industry. Finally, the five years from 2008-2012 provide hindsight to Hakes & Sauer, and reflect the pitching dominant era into which baseball has been thrown since improved testing procedures and stricter penalties for performance-enhancing drugs were introduced.

We conclude from these data that although there is clearly some variation in the value of these coefficients, the relationships among them remain relatively stable. There seems to be little evidence of significant structural shifts over this time period.[4] This makes sense, since while the game of baseball changes continuously due to the revolving door of player talent, and small rule changes, there is little reason to believe that the elements that lead to runs being scored and games being won change dramatically, especially over relatively short periods of time. Indeed, this was the conclusion that Hakes & Sauer reached in 2007. In Figure 1, we plot the value of each of the three coefficients as a time series after running the regression on each individual season.

¡¡INSERT FIGURE 1 HERE¿¿

4. Modeling Labor Market Valuation. Now that we have a quantitative understanding of how Eye, Bat, and $Power$ translate into winning, it remains to assess how individual players are compensated for these skills on the labor market. As we discussed above, baseball's labor market is complicated by its institutional constraints; thus, our task is harder, and we are less successful at it. Unlike in the previous case, where our units of observation were teams, we now build a data set of individual players. Following Hakes & Sauer, we use the natural logarithm of player salary in season t as our dependent variable, and build a model for that as a function of that player's Eye, Bat, and $Power$ in the previous season $(t-1)$. Only position players with a recorded salary and at least 130 plate appearances in the previous season were included.[5] There were 8,824 such players during 1985-2012. Control variables were added to the regression model for:

- TPA: how many total plate appearances did the player have? Clearly, if a player is playing regularly, he is likely to be paid more, regardless of how well he hits.
- $Catcher$: was the player primarily a catcher? Catcher is primarily a defensive position, so it stands to reason that catchers will make more money than players of corresponding offensive value.
- $Infielder$: was the player primarily a second baseman, shortstop, or third baseman? A similar, albeit weaker, argument can be made for

[4] Simple linear models for Eye, Bat, or $Power$ as a function of $Year$ yield p-values for the F-statistic of 0.007, 0.493, and 0.497, respectively. Although the downward trend in the value of Eye appears to be statistically significant by this measure, there do not appear to be statistically significant structural changes within the time divisions we investigated. We tested this by adding indicator variables for each time period, along with interaction terms between those indicators and $Year$, to a multiple regression model for each statistic as a function of $Year$. A nested F-test between these models and the simple linear regression model yielded p-values of 0.859, 0.697, and 0.127, respectively. A Chow test would similarly find little evidence of structural changes.

[5] Because the salary figures in the Lahman database are collected unofficially by volunteers, there are some players who have no recorded salary in a particular year.

infielders.
- $ArbElig$: was the player arbitration eligible?
- $FreeAgent$: was the player a free agent?
- Fixed effects for $Year$: in order to control for baseball inflation and other factors

Thus, our model for labor market valuation is:

$$\begin{aligned} ln(Salary) = \beta_0 &+ \beta_1 \cdot Eye + \beta_2 \cdot Bat + \beta_3 \cdot Power \\ &+ \beta_4 \cdot TPA + \beta_5 \cdot Catcher + \beta_6 \cdot Infielder \\ &+ \beta_7 \cdot ArbElig + \beta_8 \cdot FreeAgent + \boldsymbol{\gamma} \cdot \mathbf{Year}\,, \end{aligned}$$

where **Year** is a vector of indicator variables for each year from 1986-2012, and $\boldsymbol{\gamma}$ is a vector of the corresponding coefficients. The results from this regression are shown in Table 2. As we are using the same model as Hakes & Sauer, we attribute our slightly higher R^2 values to the more accurate contract data that we are using.

¡¡INSERT TABLE 2 HERE¿¿

Here, unlike our estimates of labor market productivity, we see dramatic changes in the way that players are compensated for certain skills on the labor market over time. In particular, we note that whereas a free agent outfielder of average ability in 2000 would expect to earn approximately \$36,000 more had his walk rate been ten points higher, a player with the same statistics could have expected approximately \$125,000 in additional salary in 2010, more than a threefold increase.[6] In Figure 2, we illustrate the changes in the coefficients for Eye, Bat, and $Power$ over time.

¡¡INSERT FIGURE 2 HERE¿¿

There appears to be greater fluctuation in the labor market valuations, as opposed to the labor market productivity, but the evidence in favor of structural shifts is weak.[7] As Hakes & Sauer noted that after reaching a relatively low point in 2001, the return to Eye begins to pick up in 2002 and

[6]Over the 1998-2002 period, the average values for TPA, Eye, Bat, and $Power$ among the 1,748 players the data set were 447, 0.091, 0.272, and 1.59, respectively. If such a player were a free agent outfielder, then his 2000 salary predicted by our model is \$2,424,780. However, if that players walk rate was 0.101, then his predicted salary becomes \$2,460,973, an increase of \$36,193. The corresponding difference for a free agent outfielder in 2010 with the same statistics would be \$124,655. Since we are using year fixed effects, this result controls for baseball inflation over the period.

[7]The p-values for the F-statistic of the simple linear model for Eye, Bat, or $Power$ as a function of $Year$ are 0.089, 0.098, and 0.062. The corresponding p-values for a nested F-test that includes in periodized interaction terms are 0.27, 0.846, 0.007. In particular, there is evidence that $Power$ was compensated more highly from 1998-2007 than in the other years.

2003, prior to the publication of *Moneyball*. The return then spikes in 2004, the year following the book's publication. They interpret these changes as evidence that *Moneyball* changed the labor market in baseball. In the longer-run context of the baseball labor market, while the rapid increase in the value of *Eye* surrounding the publication of *Moneyball* may or may not be causal, a more robust story is that the assimilation of the value of OBP by general managers has been a gradual change that has evolved over decades. During the 28-year window explored, the post-*Moneyball* spike in the valuation of *Eye* is just a blip in a long-term change.[8]

In their 2007 paper, Hakes & Sauer question whether the noted decline in the value of *Eye* from 2004 to 2006 is attributable to random variation, or evidence of a market correction among general managers who had come to see *Eye* as being overvalued by the market. Although the valuation of *Eye* was equally high in 2010, it does appear that the post-*Moneyball* valuation of *Eye* was abnormally high. Thus, while we see some evidence that the average valuation of *Eye* is higher in the post-*Moneyball* era, the immediate effect of the book's publication has been mitigated. Given the rigidities in the baseball players' market and the pile-on nature of market response, it is not surprising that the observed annual market correction process is neither gradual nor linear.

5. Elasticity Models. Since the three offensive performance variables have different scales, it is desirable to consider the elasticity of salary with respect to *Eye*, *Bat*, and *Power*, by taking logs on both sides of the regression model. This represents a departure from Hakes and Sauer, and allows us to interpret the coefficients as the percentage change in salary that is associated with a 1% increase in each of the explanatory variables. In Table 3, we provide results from this regression. We note that the salary return to *Eye* increased by 5.8 percent (from 0.138 to 0.146) between the 1985-97 and the 1998-2002 periods, but then shot up by 63.7 percent (from 0.146 to 0.239) in the immediate post-*Moneyball* period, before retreating 11.7 percent (from 0.239 to 0.211) during 2008-2012. It is also notable that the returns to *Power* and *Bat* increased over the extended period as well[9], although they did not grow as rapidly as the return to *Eye*. This result is evident in the *Eye*-to-*Bat* and *Eye*-to-*Power* coefficient ratios.

¡¡INSERT TABLE 3 HERE¿¿

Similarly, we can consider our labor market productivity model in terms

[8]This point is elaborated in Baumer and Zimbalist (2014), chapter one.

[9]This growth in the return to all skills should not be surprising given the sustained increase in average salaries between 1985 ($370,000) and 2012 ($3.2 million).

of elasticities. Since we include only salaries for position players in our labor market valuation, here we include only offense, and use team runs scored as the dependent variable. The results are shown in Table 4. We note that these coefficients show considerable long-term stability, with little evidence of meaningful change over an extended period of time. This corroborates our earlier observation that the elements of the game that translate into winning are relatively static.

¡¡INSERT TABLE 4 HERE¿¿

By considering the ratio of the elasticity of each statistic to salary over the elasticity of that statistic to runs scored, we gain an estimate of the relative return to each skill. We show these results in Table 5.

¡¡INSERT TABLE 5 HERE¿¿

It is interesting to note here that the ratio of these elasticities for batting average has remained remarkably consistent over these time periods. This suggests that however fairly batting average is being compensated on the labor market, that relationship has not changed much, if at all. Conversely, the return to *Power* has changed dramatically from sub-period to sub-period. It appears to now be only about half as high as it was during the home run boom of 1998-2002. However, it is also notable that between the 1985-1997 and 2008-2012 periods the return to power has stayed relatively constant, increasing by only 1.6 percent. Conversely, the return to *Eye* increased by 61 percent between the first and last periods. Moreover, the relative return to *Power* was three times higher than it was for *Eye* in the pre-*Moneyball* era, but in the most recent period is only slightly higher.[10] Further, the pattern of the relative return to *Eye* corresponds to the market correction process described earlier. Namely, the relative return to *Eye* increases in the late 1990s and the first three years of the 2000s, prior to the publication of *Moneyball*, accelerates sharply during 2003-2007 to the point of overcorrection, and then adjusts by falling 20 percent during 2008-2012.

6. Payroll Efficiency. Next we consider the relationship between team winning percentage and payroll, indexed as the share of league payroll.[11] Hakes and Sauer posit that with the improved metrics of the *Moneyball* era, teams are better able to discern the true productivity of the players and, therefore, the statistical relationship between team payroll and performance

[10]It is also interesting to observe that the return to power was highest during the apex of the steroid power boom. While an economist might scratch her head at this outcome (thinking that the plethora of home runs would lower the relative value of the same), it appears that the steroid power boom produced both an ownership and fan fascination with the long ball, only serving to increase its market value.

[11]Equivalently, each teams payroll divided by the league average payroll in that season.

should tighten.

Indeed, Hakes and Sauer found "the explanatory power of team payroll in predicting winning percentage has improved over time," citing as evidence the increase in R^2 values for a simple linear regression model for team winning percentage as a function of team payroll. Specifically, they found the R^2 in this payroll model for all 587 team-seasons between 1986 and 2006 to be 0.146. As noted above, the salary figures in the Lahman database are incomplete. Further, to derive team payroll data from individual player salaries would require knowledge of the entire 25-man roster, the players on the disabled and the proportion of a year that traded players were on the team.

We use official payroll data from the Labor Relations Department of Major League Baseball, and find that the correct R^2 over this time period is 0.228, a substantial increase over the figure reported by Hakes and Sauer. In Table 6, we present the results from applying a regression model to team winning percentage as a function of each team's LRD payroll.[12]

¡¡INSERT TABLE 6 HERE¿¿

More importantly, Hakes and Sauer cited an increase in the strength of the relationship between team winning percentage and payroll over time as evidence that the window for exploiting market inefficiency in the manner described in *Moneyball* may be narrowing; or, stated differently, Hakes and Sauer suggest that the new, improved performance metrics and correction of market inefficiencies since *Moneyball* has tightened the relationship between team payroll and win percentage. We find no such evidence. While it is true that the predictive power of payroll upon winning percentage is higher than it was in the 1980s and early 1990s, it has been much lower in the 2000s than it was in the late 1990s, and shows few signs of a long-term upward trend (see Figure 3).[13] An elegant way to test for structural breaks in this relationship is to model winning percentage as a function of normalized payroll, along with interaction terms for normalized payroll and an indicator variable for each one of the periods. None of the terms involving any period was statistically significant.

¡¡INSERT FIGURE 3 HERE¿¿

[12] LRD payroll stands for MLB's Labor Relations Department measurement of payroll. LRD payroll differs from the CBT (competitive balance tax) and players' association payroll in its treatment of benefits, deferred salary and discount rates.

[13] Again, a simple linear model for R^2 in the payroll model as a function of time is not statistically significant at the 5% level. The increase in R^2 between 1985-1997 and 1998-2002 may be a function of widening revenue inequality across teams and the Yankees' remarkable string of success during the latter period.

7. Conclusion. We employed data from 1985 through 2012, with periodization, to test the existence of, the timing of the recognition of, and the market's adjustment to an inefficiency in the baseball labor market (specifically, the hypothesized undervaluation of walks). We adapted the basic model proposed by Hakes and Sauer, extended the data set, integrated more accurate data on player service time, player salaries and team payrolls, and introduced a double-log model to properly correct for scale differences in the variables describing hitters' skills. Finally, we explored whether a core assertion of Hakes and Sauer – that there has been a linear trend for the R^2 between team payroll and win percentage to increase as a result of the use of better performance metrics and the disappearance or, at least, attenuation of the labor market inefficiency – holds up when more accurate data is used and the time period is extended.

Hakes and Sauer found evidence that the market inefficiency surrounding OBP described in *Moneyball* existed in the period prior to the publication of the book and then closed in the few years following its publication. As evidence, they cite growing equality in the expected salary return to certain batting skills from 2004-2006. We address this same question with a double-log model, and compare the relative return to these same batting skills.

Hakes and Sauer left open the question of how the labor market would react going forward. We find that the sharp increase in the return to on-base skills in the period surrounding *Moneyball* is properly viewed in the context of a longer-term evolution in valuation on the part of general managers. While *Moneyball* may well have been a catalyst in accelerating the incorporation of value of OBP into the baseball labor market, these changes were part of a more modest, longer-term trend. Moreover, the relative return to walking has fallen since Hakes and Sauers observation, confirming their suspicions that the market correction in the years following *Moneyball* would not be permanent.

Finally, using accurate payroll data from MLB, we find no evidence that the relationship between team payroll and winning percentage has strengthened over time. This runs counter to the observation made by Hakes and Sauer.

References.

Baumer, B. and Zimbalist, A. (2014). *The Sabermetric Revolution: Assessing the Growth of Analytics in Baseball*. University of Pennsylvania Press.

Hakes, J. K. and Sauer, R. D. (2006). An economic evaluation of the Moneyball hypothesis. *The Journal of Economic Perspectives* **20** 173–185.

Hakes, J. K. and Sauer, R. D. (2007). The Moneyball anomaly and payroll efficiency: a further investigation. *International Journal of Sport Finance* **2** 177–189.

LAHMAN, S. (2013). Sean Lahman's Baseball Database. http://www.seanlahman.com/baseball-archive/statistics/.

LEWIS, M. (2003). *Moneyball: The Art of Winning an Unfair Game*. WW Norton & Company.

SARKAR, D. (2008). *Lattice: Multivariate Data Visualization with R*. Springer, New York.

R CORE TEAM (2012). R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.

	1985-2012	1985-1997	1998-2002	2003-2007	2008-2012
Constant	0.500	0.500	0.500	0.500	0.500
Eye	1.713	1.862	1.555	1.758	1.419
	(0.081)	(0.136)	(0.167)	(0.157)	(0.217)
Power	0.264	0.272	0.219	0.260	0.319
	(0.016)	(0.027)	(0.031)	(0.034)	(0.039)
Bat	3.005	2.907	3.266	2.911	3.087
	(0.074)	(0.119)	(0.150)	(0.162)	(0.174)
$\beta(Eye)/\beta(Bat)$	0.570	0.641	0.476	0.604	0.460
$\beta(Eye)/\beta(Power)$	6.497	6.839	7.117	6.747	4.448
N	798	348	150	150	150
R^2	0.814	0.773	0.882	0.827	0.808

TABLE 1

Estimates of Labor Market Productivity. The dependent variable is Team Winning Percentage. Note: In all tables, standard errors are in parentheses below the coefficients.

	1985-2012	1985-1997	1998-2002	2003-2007	2008-2012
Eye	2.000	1.667	1.434	2.827	2.782
	(0.210)	(0.287)	(0.417)	(0.521)	(0.562)
Power	0.667	0.611	0.745	0.731	0.578
	(0.034)	(0.049)	(0.070)	(0.086)	(0.086)
Bat	3.672	3.402	3.910	4.274	3.828
	(0.235)	(0.334)	(0.473)	(0.614)	(0.579)
$\beta(Eye)/\beta(Power)$	3.000	2.727	1.926	3.869	4.815
$\beta(Eye)/\beta(Bat)$	0.545	0.490	0.367	0.614	0.727
N	8824	3668	1748	1707	1701
R^2	0.771	0.735	0.787	0.722	0.719

TABLE 2

Estimates of Labor Market Valuations. The dependent variable is $\ln(Salary)$.

	1985-2012	1985-1997	1998-2002	2003-2007	2008-2012
Eye	0.167	0.138	0.146	0.239	0.211
	(0.019)	(0.025)	(0.038)	(0.048)	(0.050)
Power	1.050	0.920	1.178	1.213	0.978
	(0.054)	(0.075)	(0.111)	(0.137)	(0.135)
Bat	0.904	0.826	1.009	1.121	0.894
	(0.062)	(0.087)	(0.126)	(0.165)	(0.150)
$\beta(Eye)/\beta(Bat)$	0.185	0.167	0.144	0.213	0.236
$\beta(Eye)/B\beta(Power)$	0.159	0.150	0.124	0.197	0.216
N	8824	3668	1748	1707	1701
R^2	0.769	0.733	0.786	0.720	0.717

TABLE 3

Estimated Labor Market Valuation Elasticities. The dependent variable is $\ln(Salary)$.

	1985-2012	1985-1997	1998-2002	2003-2007	2008-2012
ln(*Eye*)	0.277	0.267	0.305	0.230	0.255
	(0.012)	(0.020)	(0.026)	(0.024)	(0.026)
ln(*Power*)	0.854	0.982	0.774	0.957	1.027
	(0.029)	(0.048)	(0.075)	(0.066)	(0.062)
ln(*Bat*)	1.758	1.671	1.850	1.960	1.69
	(0.029)	(0.047)	(0.062)	(0.064)	(0.57)
$\beta(Eye)/\beta(Bat)$	0.158	0.160	0.166	0.117	0.151
$\beta(Eye)/\beta(Power)$	0.324	0.272	0.394	0.240	0.248
N	798	348	150	150	150
R^2	0.911	0.908	0.918	0.910	0.917

TABLE 4

Estimated Labor Market Productivity Elasticities. The dependent variable is ln(*Team Runs Scored*).

	1985-2012	1985-1997	1998-2002	2003-2007	2008-2012
Eye	0.602	0.514	0.478	1.040	0.827
	(0.072)	(0.102)	(0.131)	(0.236)	(0.213)
Power	1.230	0.937	1.522	1.267	0.952
	(0.075)	(0.089)	(0.206)	(0.167)	(0.144)
Bat	0.514	0.495	0.545	0.572	0.530
	(0.036)	(0.054)	(0.071)	(0.086)	(0.091)

TABLE 5
Relative Return to Offensive Skills

	1985-2012	1985-1997	1998-2002	2003-2007	2008-2012
Payroll	0.089	0.090	0.107	0.085	0.074
	(0.006)	(0.011)	(0.013)	(0.012)	(0.013)
N	798	348	150	150	150
R^2	0.214	0.166	0.304	0.257	0.171

TABLE 6

Payroll Efficiency. The dependent variable is Team Winning Percentage.

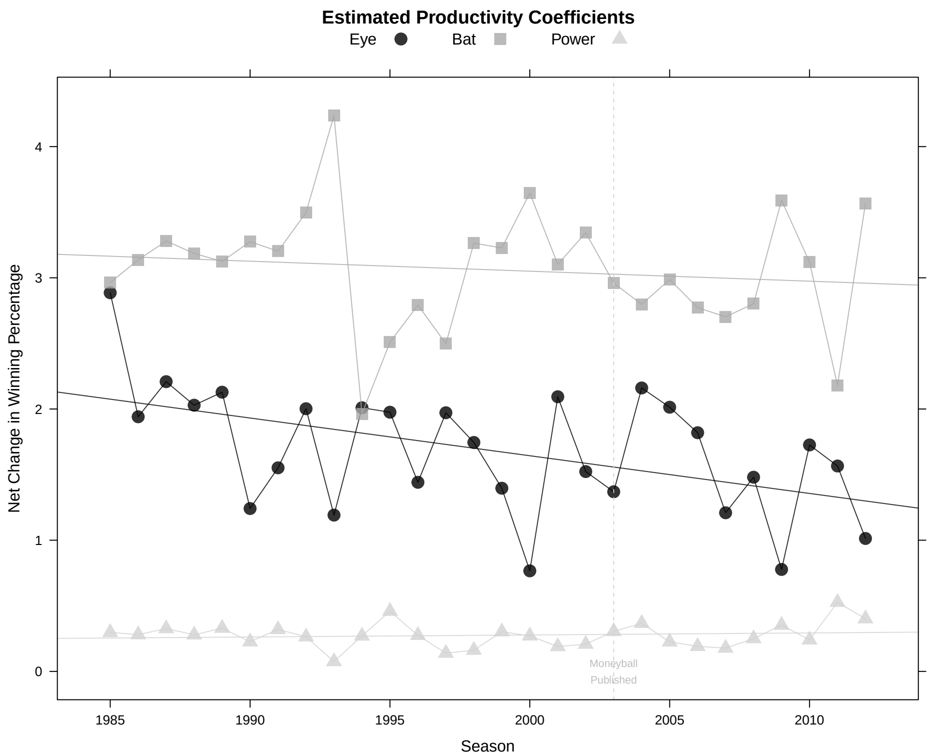

FIG 1. *Estimated Labor Market Productivity Coefficients*

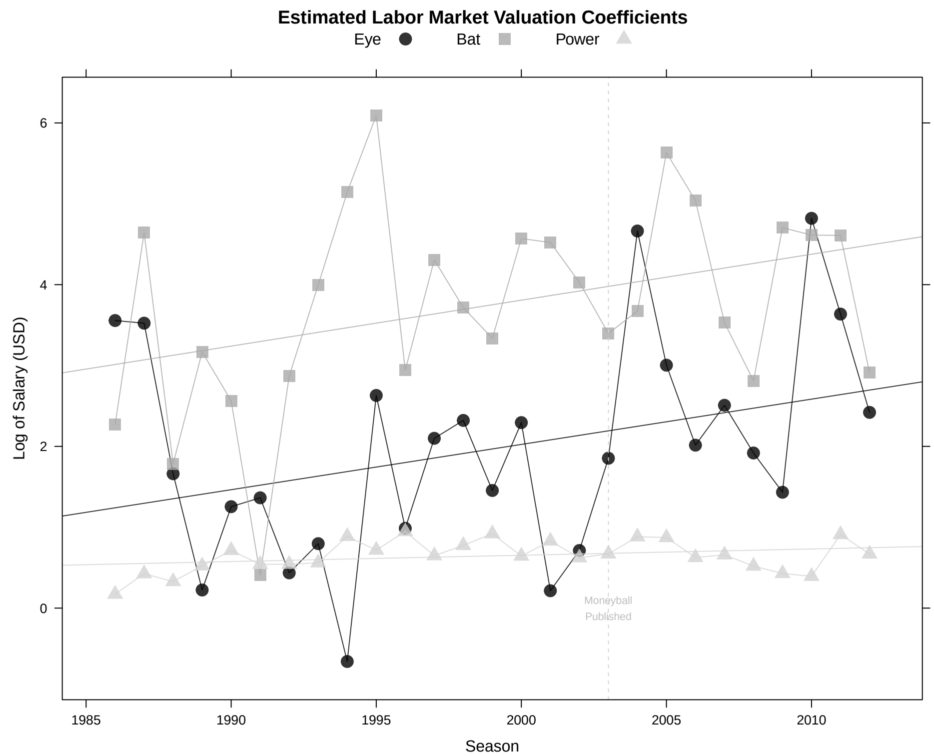

Fig 2. *Estimated Labor Market Valuation Coefficients*

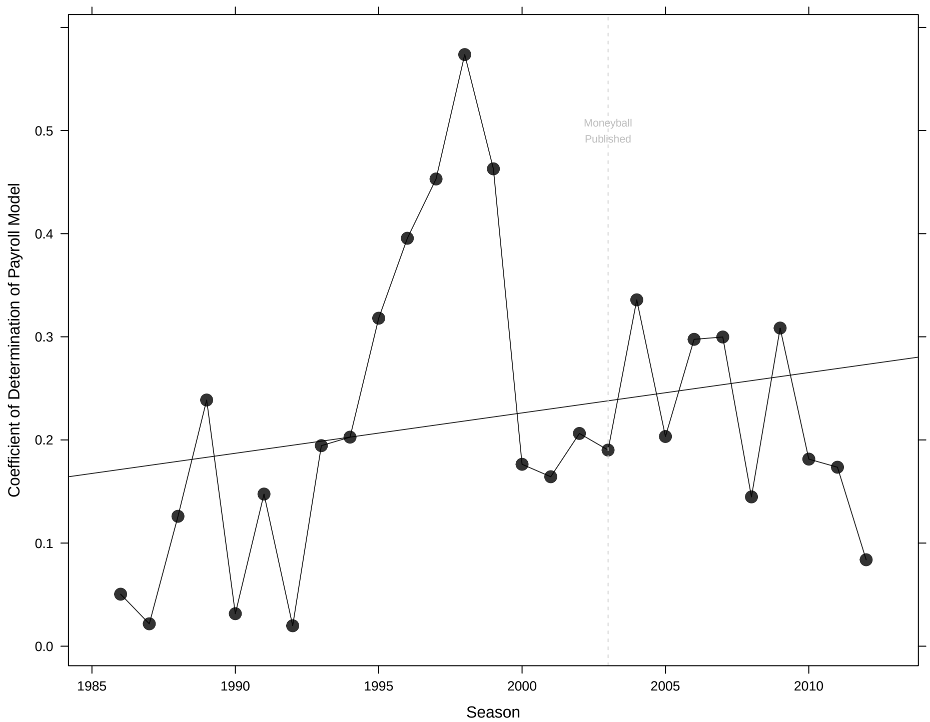

FIG 3. *Coefficient of Determination in Payroll Model over Time - Each dot represents the value of the coefficient of determination (R^2) in a simple linear model for team winning percentage as a function of indexed payroll in a given season.*

DEPARTMENT OF MATHEMATICS & STATISTICS
E-MAIL: bbaumer@smith.edu

DEPARTMENT OF ECONOMICS
E-MAIL: azimbali@smith.edu