Journal of Sports Economics

http://jse.sagepub.com

Working in the Land of the Metricians

David J. Berri and John Charles Bradbury

Journal of Sports Economics 2010; 11; 29
DOI: 10.1177/1527002509354891

The online version of this article can be found at:
http://jse.sagepub.com/cgi/content/abstract/11/1/29

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
The North American Association of Sports Economists

Journal of Sports Economics
11(1) 29-47

DOI: 10.1177/1527002509354891
http://jse.sagepub.com
SAGE

Working in the Land of the Metricians

David J. Berri,[1] and John Charles Bradbury[2]

Abstract

The research of sports economists often addresses issues of interest to nonacademics. The shared interests often lead to interactions that have benefits and costs. The benefits center on nonacademic research—found in the "sabermetric" and "APBRmetric" communities—that can inform the work of academic economists studying sports. However, nonacademic research should be interpreted with caution because it is not subject to an academic peer review. This essay discusses how economists can benefit from sabermetric advances while avoiding its pitfalls.

Keywords

sports economics, research methods, baseball, basketball, sabermetrics

Introduction

Across the last decade, the field of sports economics has expanded from an occasional published paper to a field with its own journals and organizations. Economists are attracted to this field partly because it is blessed with an abundance of data on worker productivity. As Lawrence Kahn (2000, p. 75) observes, "There is no research setting other than sports where we know the name, face, and life history of every production worker and supervisor in the industry."

Another facet of sports is equally alluring: many economists—who were probably sports fans before they became economists—simply find sports to be extremely

[1] Department of Economics and Finance, Southern Utah University, Cedar City, Utah
[2] Department of Health, Physical Education, and Sport Science, Kennesaw State University, Kennesaw, Georgia

Corresponding Author:
David J. Berri, Department of Economics and Finance, Southern Utah University, 351 West University Boulevard, Cedar City, UT 84720.
Email: berri@suu.edu

interesting. This second reason is often downplayed by sports economists; after all, nonsports fans often sit in judgment of this research. If the research findings are not relevant to economists in general, the academic community will ignore them. As Jewell (2006, p. 10) recently stated, "It is incumbent upon sports economists to explain why their research can be generalized beyond the sports industry, something at which they have become increasingly adept."

While sports economists are making their work interesting to economists who do not follow sports, another population of people has taken an interest in sports economics research. This second population consists of individuals intensely interested in sports but not familiar with economic research methods and practices. This population of nonacademics introduces an issue for sports economists that other economists working on nonsports research do not generally face. Specifically, how should economists react to the work and ideas of the nonacademic analysts who are often seeking to answer the same questions?

Perhaps the most famous nonacademic sports analyst is Bill James. In the late 1970s and early 1980s, James popularized his unique approach—which he dubbed "sabermetrics" in homage to SABR, the Society for American Baseball Research—to baseball analysis that involved objective empirical analysis. James's *Baseball Abstracts* generated a legion of fans that would use James's approach to analyzing baseball. Thus, sabermetrics became a research movement. It was not long before a similar movement—dubbed APBRmetrics in honor of the Association for Professional Basketball Research (APBR)—developed in basketball. We refer to these as the "metricians" to get both the "saber" and APBR variety under a single name. Fundamental to the rest of the article, we observe that these communities of nonacademic sports analysts evolved outside the typical parameters of an expert peer-review system familiar to most academics.

The interaction between these laymen and academics has not always struck a positive tone. Some metricians accuse the academic group of intellectual snobbery—that is, deliberately snubbing valid research only because the authors lack credentials. Phil Birnbaum, a frequent critic of both of the authors of this article, summarizes this complaint (Birnbaum, 2006),

> Years of sabermetric advances, as valid as anything in the journals, are dismissed out of hand because of a kind of academic credentialism, an assumption that only formal academic treatment entitles a body of knowledge to be considered, and the presumption that only the kinds of methods that econometricians use are worthy of acknowledgement.
>
> The truth is, there's pretty decent statistical rigor in some of what us amateurs have done. In "Solving DIPS," [a measurement debate] a bunch of really smart people, statistically literate, probably no less intelligent than academic economists and much better versed in sabermetrics, do an awesome and groundbreaking job of determining the causes of variation in the results of balls in play. [Bracketed item added by the authors for clarity.]

Birnbaum (2006) considers sabermetricians to be "no less intelligent than academic economists" and superior to economists in their understanding of baseball. This

statement reveals a curious worldview. On one hand, the aspect that is universal across both groups—members of both communities have been devoted sports fans since an early age—is considered unique to the nonacademic sports analysts. On the other hand, when it comes to the aspect that is unique to academics—academia normally involves many years of advanced training and requires its participants to be judged competent by their peers in a "publish or perish" environment—metricians demand equal recognition. In our view, this mentality begets misplaced confidence.

As "awesome and groundbreaking" as Birnbaum (2006) perceives sabermetric work to be, researchers seeking to study sports phenomena should proceed with caution when using informally vetted metrician findings. Often the merit of ideas offered by this community is judged by a consensus of pseudonymous avatars, many of whom appear to lack training and experience with advanced research methods. Consequently, much of what passes for research among metricians would not pass through the academic peer-review process of any academic discipline. Therefore, a full-scale reevaluation and formal demonstration that these techniques are sufficiently rigorous for use should be required by journal referees and editors.

That being said, researchers who validate metricians' findings should acknowledge the origins of their approach. Although we are critical of the hobbyist community's informality, we must acknowledge that metricians are responsible for many important discoveries that should inform sports economists in their research. In fact, we have observed instances where economists have made elementary mistakes in understanding sports that would not have been made if they had paid attention to metricians' findings. It is our hope that academic researchers using sports as their laboratory will review the analysis from all relevant outlets, including the metrician community.

With these thoughts in mind, we now proceed to three specific questions:

1. What can the sports economist learn from the metrician research?
2. What issues arise when sports economists review the metrician research?
3. How should sports economists handle the less savory interaction aspects when we journey into the land of the metricians?

Sports Economics and the SABRmetricians

Simon Rottenberg wrote a paper widely believed to be the first sports economics paper[1]—"The Baseball Player's Labor Market"— in 1956. The sport of choice for Rottenberg—Major League Baseball—was the same sport chosen by researchers following Rottenberg's footsteps. Perhaps the most famous of these followers is Gerald Scully. In 1974, Scully published "Pay and Performance in Major League Baseball," which offered the first empirical estimate of workers' marginal revenue products (MRPs). Scully's method is well known to sports economists. His approach

involved two steps. First, he quantified the impact of player performance on winning. He then linked wins to team revenue.

Scully's first step requires understanding how player performance translates into on-field performance (a subject heavily studied by sabermetricans). As proxies for player performance, Scully used slugging percentage (total bases divided by at-bats) for hitters and strikeout-to-walk ratio for pitchers. Many economists continue to use slugging average to measure hitter quality,[2] but earned run average (ERA) has been the more popular choice to measure pitchers. Although both slugging average and ERA have proven to be popular choices, neither is a particularly "good" measure of a baseball player's productivity.

Bradbury (2008) sets two criteria for evaluating sports metrics as proxies for player performance: (a) the accuracy to which the metric measures value—in this case, how the metric is correlated with run production for hitters and run prevention for pitchers—and (b) the degree to which the metric reflects skill. From these criteria, we demonstrate what it means to have a "better" metric, starting with hitters. The primary objective of a hitter is to score runs. Consequently, a performance metric needs to be evaluated relative to this objective. The following equation illustrates this idea:

$$\text{Runs scored per game } = a_0 + a_1 * \text{performance measure } + e_{1t} \qquad (1)$$

For example, consider batting average (hits divided by at-bats) as a performance measure. This metric can be traced back to the work of H. A. Dobson, in 1872, and remains the most commonly cited measure of a hitter's performance. Using a sample of team data to predict runs scored, batting average as the lone performance measure in Equation 1 explains approximately 65% of the variation in runs scored per game.[3] Relative to slugging percentage—which explains 78% of the variation in runs scored per game—batting average is clearly an inferior metric.[4] One can do even better, though, with on-base percentage plus slugging percentage (OPS). This metric—popularized by Thorn and Palmer (1984)—can explain 89% of the variation in runs scored. The calculation and explanation of OPS is not much beyond what was seen for slugging percentage; and because OPS has a better connection to outcomes, sports economists have recently turned to this measure.[5]

Once again, we return to the word "better." So far, we have grounded this term in the concept of explanatory power. However, simplicity is also important. Performance measures are generally used to tell a larger story about economics. Because researchers do not have unlimited space in an academic article, increasing the words devoted to explaining a performance measure can ultimately limit what is said about the bigger story. Consequently, sports economists prefer a parsimonious metric to a more complex measure. This means that sports economist should be willing to embrace OPS, because its explanatory power trumps what we see from slugging percentage. Using more complex metrics, though, may yield little advantage.

To illustrate, consider the work of Blass (1992). Blass estimated the following equation—which he did not create himself—to explain runs scored per game.

$$\begin{aligned}\text{Runs scored per game} &= b_0 + b_1 * 1\text{B} + b_2 * 2\text{B} \\ &+ b_3 * 3\text{B} + b_4 * \text{HR} + b_5 * \text{NBB} + b_6 * \text{HBP} + b_7 \\ &* \text{SB} + b_8 * (\text{GDIP} + \text{CS}) + b_9 * \text{SF} + b_{10} * \text{OUTS} + e_{2t},\end{aligned} \tag{2}$$

where 1B = singles per game; 2B = doubles per game; 3B = triples per game; HR = home runs per game; NBB = nonintentional walks per game; HBP = hit by pitch; SB = stolen bases per game; GIDP = ground into double plays per game; CS = caught stealing per game; SF = sacrifice flies per game; and OUTS = outs per game.

The linear regression estimated by Blass explains 93% of the variation in runs scored per game. Blass's estimate explains more variance than OPS. However, the gains come at a cost of increased complexity; and it is possible that many will conclude the cost of increased complexity trumps the small benefits gained from increased explanatory power.

Let us say one decided that the benefit of increased explanatory power did exceed the cost of additional complexity. If one made such a choice, a lesson from the sabermetric community would prove relevant. Specifically, there is a potential bias in the regression estimates offered by Blass.

Using the relationship between stolen bases and foot-speed, Turocy (2005) highlights the problem of purely regression-based run estimators in baseball. According to Turocy, though the added explanatory power of a regression is superior to some simple metrics, the coefficient weights suffer from omitted variable bias due to the correlation of included metrics with other unmeasured important aspects of the game. "Linear weights" is therefore a superior estimator of run production, because it does not suffer from omitted variable bias.

This metric was originally developed by operations research analyst Lindsey (1963) and updated by sabermetricians Thorn and Palmer (1984). Linear weights estimates contributions to baseball events using play-by-play data to weight the run-generating probabilities for individual events; thus, its estimates avoid the omitted variable bias problem. Albert and Bennett (2003) demonstrated that linear regression and linear weights methods yield similar results with the difference being that linear regression (p. 203):

> discovers *overall tendencies* in run production . . . .This may be good or bad. It's bad in the sense that we are not sure exactly what each measure represents . . . . It's good that the [linear regression] model may capture aspects of baseball within the data that are not measured explicitly.

In this case, we have an academic approach (Lindsey) refined by metricians (Thorn and Palmer, 1984), with further contributions made by academics (Albert & Bennett,

2003; Turocy, 2005). The interplay between these groups has resulted in a metric that is accurate and correctly specified. One still might wonder whether the benefit of using this more complex approach could be justified by the cost. Nevertheless, the story does highlight how academics can learn from metricians.

A similar story of the interaction between sports economists and the metricians can be seen when we turn to the evaluation of pitchers, which highlights the importance of Bradbury's second criterion for examining performance measures. Scully (1974, 1989) used strikeout-to-walk ratio to measure pitcher performance; however, Zimbalist (1992b)[6] and Krautmann (1999) argue that ERA is a better measure of pitcher quality because ERA has greater explanatory power. However, one also has to consider consistency of performance when choosing performance measures. If a measure varies considerably for an individual player over time, it is likely that the metric is heavily polluted by luck and thus reflects irrelevant information that should not be used for valuing skill. Scully appeared to understand this point with his choice of the strikeout-to-walk ratio to measure pitcher quality.

Later research from outside academia confirmed Scully's original approach and suggested including home-run prevention when evaluating pitchers. On the popular sabermetric Web site *Baseball Prospectus*, McCracken (2001) published an essay that suggested pitchers have little control over hits allowed on balls hit in play. The lack of control occurs because balls hit into play are heavily influenced by defensive performance and random chance. As a consequence, the evaluation of pitchers should avoid measures—such as ERA and WHIP—that depend on the players surrounding the pitcher. McCracken suggested using three defense-independent pitching statistics (DIPS) to evaluate pitchers (that are not biased by the luck and fielding contributions on balls in play): strikeouts, walks, and home runs. This novel approach judges pitcher performances when no fielders are involved; thus, the information revealed tells us quite a bit more about the performance of the pitcher than ERA or WHIP.[7]

If we focus on explanatory power, we see that a team's ERA—relative to the DIPS factors—is much better at explaining the number of runs a team allows. However, explanatory power is not the only criterion on which we should evaluate performance metrics. ERA and runs allowed are virtually the same, and their high correlation should be expected. ERA is just a description of what happened, and that description includes factors beyond a pitcher's control—luck and fielder contributions—that should not be included when evaluating pitchers.

When we turn to consistency, we see that ERA is not a particularly good measure of performance. Again, true talent should persist over time. As Table 1 indicates, among pitching statistics, strikeouts, walks, and home runs are strongly correlated from season to season for pitchers, while batting average on balls in play and ERA are not. This supports McCracken's assertion that pitchers have little control over hits on balls in play. While ERA scores well on the first criterion, it fails the second requirement miserably.

Table 1. The Consistency of Hitters and Pitchers in Baseball, 1980 to 2005

	r	r^2
Statistic for hitter		
Batting average	.47	.22
On-base percentage (OBP)	.64	.41
Slugging average (SLG)	.67	.45
On-base percentage plus slugging average (OPS)	.65	.43
Linear weights	.70	.49
Statistic for pitcher		
Batting average on balls put in play (BABIP)	.24	.06
Earned run average (ERA)	.37	.14
Home runs per nine innings	.47	.19
Walks per nine innings	.64	.42
Strike-outs per nine innings	.79	.62

Source: Bradbury (2008)

Of course, a metric that captures the independent contribution of pitchers as opposed to one that is clouded by other defense and luck is important to sports economists. If one were to evaluate the labor market for pitchers using ERA, WHIP, or any other pitching metric heavily influenced by performance on balls in play, it would likely value pitchers improperly.[8] Consequently, we see Scully's general approach is confirmed by a metrician, demonstrating what the nonacademic sports research community can contribute to sports economics research.

Sports Economics and the APBRmetricians

Our discussion of baseball suggests economists can learn much from "walking with the metricians." And in terms of lowering the cost of research, it is certainly tempting to use metrics developed by others. However, one has to use caution. To see this, we now consider a few of the metricians' measures of a basketball player's performance.

We begin with the "efficiency" metrics, National Basketball Association (NBA) efficiency, and John Hollinger's player efficiency rating (PER). The former is just a matter of arithmetic.

$$\begin{aligned}\text{NBA Efficiency} &= \text{PTS} + \text{ORB} + \text{DRB} + \text{STL} + \text{BLK} \\ &+ \text{AST} - \text{TO} - \text{MSFG} - \text{MSFT},\end{aligned} \tag{3}$$

where PTS = points scored; ORB = offensive rebounds; DRB = defensive rebounds; STL = steals; BLK = blocked shots; AST = assists; TO = turnovers; MSFG = missed field goals; and MSFT = missed free throws.

NBA Efficiency is perhaps the oldest summary measure in basketball. It is quite similar to the TENDEX model developed first developed by Dave Heeran in 1959.[9] Robert Bellotti's Points Created model, published in 1988, is also very similar.[10]

PER is a more complex measure created by Hollinger (2002). To see the complexity of PER, consider the simplified version of PER that Hollinger labels Game Score.

$$\begin{aligned}\text{Game Score} &= \text{PTS} + 0.4 * \text{FGM} - 0.7 * \text{FGA} - 0.4 * \text{MSFT} + 0.7 * \text{ORB} \\ &+ 0.3 * \text{DRB} + \text{STL} + 0.7 * \text{AST} + 0.7 * \text{BLK} - 0.4 * \text{PF} - \text{TO},\end{aligned} \tag{4}$$

where FGM = field goals made; FGA = field goals attempted; and PF = personal fouls.

In addition to the elements of Game Score, PER adjusts for both game pace and minutes played.[11] These PER adjustments, which make the calculations quite a bit more complicated than what we see in Equation 4, do not have much impact on the actual evaluation of players. For example, we have a sample of 896 player observations from the 2007-08 and 2008-09 seasons. Across this sample, PER—which is a per minute measure—and Game Score per 48 min have a .99 correlation. We see a similar story for Game Score and NBA Efficiency. Once again, these measures look quite different but there is a .99 correlation between NBA Efficiency and Game Score across these 2 years.

With closer scrutiny, we see that the similar story each metric tells us about scoring is the key reason for the high levels of correlation observed. Consider a player who takes 24 shots, 12 from three-point range and 12 from two-point range. If this player makes three from beyond the arc and four from within the arc, he will have scored 17 points and missed 17 field goals. So his NBA Efficiency will be unchanged. This tells us that if a player converts more than 25% of his shots from three-point range and more than 33% from two-point range, the more he shoots the higher will be his NBA Efficiency. There were 325 players in our aforementioned sample who took at least 100 shots from two-point range in 2008-09. Of these, only one player failed to hit 33% of these shots. From three-point range, there were 158 players who took at least 100 shots and none shot worse than 25%.

When we turn to Game Score, we see even lower break-even points. Game Score subtracts 0.7 for each field goal attempt. So if a player takes 10 shots, he will see his Game Score fall by 7. To overcome this loss, though, he only has to make 2.9 shots from two-point range.[12] And no player who took at least 100 two-point field goal attempts in 2008-09 failed to convert on 29% of these shots. If we look at three-point shots, the break-even point is only 20.1% and, again, all players who attempted 100 such shots surpassed this threshold.

These low break-even points tell us that the more an NBA player shoots, the higher will be his NBA Efficiency or Game Score (or PERs). However, while metricians may find the measures useful, sports economists are also looking beyond just performance to the marginal contribution to winning and, then, on to the

determination of player economic value. From the economist's perspective, inefficient scoring does not help a team win.

To see this argument, consider the link between wins and the performance metrics just described:

$$\text{Team winning percentage } = c_0 + c_1 * \text{performance measure} + e_{3t} \qquad (5)$$

Looking at team data in our same sample from 1987-88 to 2007-08—or 591 team observations—we see that a team's NBA Efficiency per game explains 32% of the variation in team winning percentage. PERs and Game Score per game explains only 33% and 31%, respectively.[13] It is interesting to note that Basketball-Reference.com—the premier site for basketball statistics—does not report NBA Efficiency for players. It does refer to PER, though, as an "advanced" measure. It appears that the online community actually prefers complexity, and PER is more difficult to calculate than NBA Efficiency. However, this complexity, though, does not provide any additional explanatory power.

To put these results in further perspective, consider Wins Produced, a measure reported by Berri (2008).[14] Wins Produced explains 94% of team wins. Wins Produced considers the same box score statistics, but break-even shooting percentages according to this metric are 33% for three-point range and 50% from two-point range. One should note, though, that in calculating Wins Produced, a team-defensive adjustment is used, which takes into account the number of points scored by the opponent, opponent's field goals made, opponent's free throws made, opponent's turnovers (that are not steals), and team rebounds that change possession. If you add this team-defensive adjustment to Equation 5, you can increase explanatory power for NBA Efficiency to 58%. For PERs and Game Score, explanatory power rises to 56% and 60%, respectively.[15]

So the metrics that use "efficiency" in their names are actually of limited usefulness. Ironically, this is because these models reward inefficiency. Consequently, these models are not well connected to winning and, therefore, are not useful in determining the economic value of players.

Even without the benefits of regression analysis, it seems likely that people can see the problems with the "efficiency" metrics. For example, Allen Iverson—an inefficient scorer—often is ranked very high by these efficiency metrics. However, when Iverson left the Philadelphia 76ers in 2006, the Sixers suddenly improved. And when Iverson left the Denver Nuggets in 2008, the Nuggets also improved.

Such results have led people to conclude that the box score statistics, which are used to construct the "efficiency" metrics, must be flawed. And that has led people to consider the plus–minus measure. Plus–minus has historically been used in hockey. This statistic is "calculated by subtracting the total number of goals allowed by a player's team while the player is on the ice (at even strength or on the power play) from the total number of goals scored by the player's team while the player is on the ice (at even strength or short-handed)."[16]

The purpose of tracking statistics is to separate a player from his teammates. The plus–minus measure, though, provides a player evaluation that fails to accomplish this objective. To see this, let us consider Bradbury's second criteria, consistency across time. Regressing a player's plus–minus this season on his plus–minus last season reveals that the latter explains only 9% of the variation in the former.[17] When we complete a similar exercise for other statistics tracked for hockey players—such as shooting percentage, assists, goals, points, penalty minutes, and shots on goal—we see a level of explanatory power that ranges[18] from 39% to 80%. In sum, every other hockey statistic tracked for skaters exhibits a far higher level of consistency across time. And this suggests that relative to plus–minus, every other statistic captures more of the player's individual skill and less the happenstance of his teammates' identity.[19]

So, it appears that plus–minus is not a particularly powerful metric in hockey. Unfortunately, when we turn to basketball, we see a very similar story. Plus–minus in basketball—just as we saw in hockey—measures how well a team does with and without a player on the court. And just as we saw in hockey, plus–minus on the court is quite inconsistent across time. Looking at data from 2006-07 to 2008-09, we see that only 23% of a player's plus–minus in a current season is explained by his plus–minus the previous campaign.[20]

Again, the problem with plus–minus is that a player's teammates can affect his measure. Consequently, as a player's teammates change, a player will see his plus-minus value fluctuate. To combat this problem, people have introduced adjusted plus–minus. The approach involves using a regression that is designed to control for the impact of a player's teammates on his plus–minus value.[21] Although in theory the quality of teammates has been controlled, two issues suggest otherwise.

First is the issue of statistical significance. Table 2 presents the adjusted plus–minus results for the 10 players who logged at least 1,000 min for the LA Lakers in 2008-09. For only two players—Lamar Odom and Jordan Farmer—do we see a coefficient that is at least twice the value of the standard error. Only one other player—Derek Fisher—has a coefficient that is 1.5 times the value of the standard error. In sum, by traditional measures of statistical significance, most players on the Lakers—according to this metric—did not have a statistically significant impact on team outcomes in 2008-09.

What we see for the Lakers is what we tend to see for all players. Looking at 666 player observations from the 2007-08 and 2008-09 seasons, we find that only 10.2% of players evaluated had an adjusted plus–minus coefficient that was at least twice the value of the corresponding standard error. And only 20.4% of coefficients were at least 1.5 times the value of the standard errors. In sum, for most players, it appears the results are not statistically significant.

The authors of this method have argued that increasing the amount of data results in smaller standard errors. And that is true. Coefficients were estimated for 292 players who played in both 2007-08 and 2008-09. For this data set, 14.7% of players had

Table 2. Adjusted Plus–Minus Values for the LA Lakers in 2008-09

Player	One-Year Adjusted Plus–Minus	Standard Error
Lamar Odom	16.64*	4.43
Kobe Bryant	6.69	6.11
Pau Gasol	6.04	5.51
Andrew Bynum	3.78	4.78
Luke Walton	1.14	5.55
Sasha Vujacic	0.61	5.31
Vladimir Radmanovic	−1.24	4.37
Trevor Ariza	−2.78	5.16
Derek Fisher	−11.09**	6.95
Jordan Farmar	−16.99*	6.73

* Coefficient is at least twice the value of the standard error.
** Coefficient is at least 1.5 times the value of the standard error.

a coefficient that was twice the value of a standard error. Looking at the 1.5 threshold, we find that 26.0% of coefficients surpass this mark.

An even greater gain is seen if 5 years of player data are examined. Estimating a coefficient for 373 players who played for five seasons, we see that 38.9% of coefficients are at least twice the value of the standard error. And 50.4% surpass the 1.5 threshold. Although more data do increase the level of statistical significance, it is still the case that most players—even when 5 years of data are used—are not found by this method to have a statistically significant impact on outcomes.

The issue of statistical significance is not the only problem with adjusted plus–minus. The issue of consistency over time—which we discussed with respect to the unadjusted plus–minus metric—remains a problem. Looking at 239 players who played in both 2007-08 and 2008-09, we see that only 7% of a player's adjusted plus–minus in the latter year is explained by what he did in the prior campaign. So it does not appear that the adjusted plus–minus results provide much information about a player's future performance.[22]

So, for the sports economist, despite quite a bit of support among the metricians, efficiency measures and plus–minus are problematic. PER—despite appearances—is not significantly different from the simplistic NBA Efficiency model. Both models overvalue inefficient scoring and do not do a very good job of explaining current wins. Turning to plus-minus, how many points are scored and surrendered when a player is on the court should be linked to current wins. However, efforts to separate the player from his teammates have not proven successful. For most players, the adjusted plus–minus approach argues that the player did not have a statistically significant impact on outcomes. Second, adjusted plus–minus cannot predict the future very well. Consequently, despite the popularity of these approaches among the metricians, neither is an improvement over what has been published in academic journals.[23]

These approaches, though, do provide a cautionary tale for sports economists seeking assistance from the metricians. The story of Linear Weight and DIPS highlights how metricians can help sports economists. Our examination of PER and adjusted plus–minus, though, suggests economists should be careful. The metricians do not provide the same peer review that academic journals do. Consequently, just because a work is accepted among the metricians, it does not necessarily mean this work would pass muster in the academic community.

Sports Economics in the Blogosphere

A growing trend in the transfer of ideas among academics is the publication of regularly updated Weblogs, more commonly known as "blogs." *The Semi-Daily Journal of Economist Brad DeLong* (delong.typepad.com), *Greg Mankiw's Blog* (gregmankiw.blogspot.com), *Freakonomics* (freakonomics.blogs.nytimes.com), *The Conscience of a Liberal* (krugman.blogs.nytimes.com), and *Marginal Revolution* (marginalrevolution.com) are a few examples of outlets where economists present many of their ideas to a general audience using the Internet. Because of the wide popularity of sports, it is no surprise that there are several blogs devoted to sports economics. A sample would include

The Sports Economist (thesportseconomist.com)
Sabernomics (sabernomics.com)
The Wages of Wins Journal (dberri.wordpress.com)
Hawkonomics (hawkonomics.blogspot.com)
International Journal of Sport Finance Blog (ijsf.wordpress.com)

In addition to disseminating ideas more widely, blogs offer the opportunity to discuss and debate with the audience. Usually, blog authors focus on current events and items of general interest that will appeal to an audience beyond academics; yet, often progressive dialogue takes place in these forums. However, no matter the pedigree of the participants or the overall quality of ideas, these forums are no substitute for the formal peer-review process that typically governs economics research.

Even after academics gain notoriety for their academic work, online commentary is not the same as academic research, even if the commentary is offered by an already-accepted reputational brand. For example, Nobel Prize winning economist Paul Krugman writes a regular column and blog for *The New York Times*; however, he does not present groundbreaking research in this forum and expect economist colleagues to address his commentary with the same level as he does his academic papers. Regardless of the usefulness of commentary, blogs should be considered a source of ideas and inspiration for further research rather than a serious research outlet. Good ideas developed on blogs should be written up, tested rigorously, and then published through normal channels.

This fact is sometimes lost to participants unfamiliar with economics research methods and evaluation protocol. This is where the Kruger-Dunning effect can sometimes manifest itself among participants. Kruger and Dunning (1999) find that many nonexperts tend to overestimate their abilities. "Not only do these people reach erroneous conclusions and make unfortunate choices, but their incompetence robs them of the metacognitive ability to realize it."

The knowledge disparity between academics and interested laymen sometimes leads to unpleasant discourse. Cyberspace is an equalizer in presenting ideas; and, because there is no barrier to publication, some participants often confuse ease of access in the same forum as equal expertise. When nonacademics discover their ideas are not being accepted, unpleasant behavior can result. Examples of "unpleasant behavior" include commentators leaving essentially the same comment over and over again, often under different aliases; and/or comments can devolve into personal attacks. The anonymous nature of commentators can often make misbehavior easier. As a consequence, some blog owners have gone so far as to remove the comments option to avoid the unpleasantness. And even when comments are allowed, the more popular blogs find it necessary to monitor the comments offered by readers.

Despite the behavior of commentators frequently referred to as "trolls," the online interaction has proven to be net beneficial. Most commentators are pleasant and offer useful insight. Consequently, very few (if any) academics ever walk away from the blogging experience once they have developed an audience. Apparently, by their own behavior, sports economists who are aware of the costs of blogging can certainly reap many benefits; and thus, blogging can ultimately improve the research generated by sports economists.

Notes

1. Peach (2004) observed that Veblen (1899) offered thoughts on the economics of sports. So one could argue the field of sports economics begins in the 19th century.
2. Sports economics studies using slugging percentage include Sommers and Quinton (1982), Raimondo (1983), Sommers (1990), Bruggink and Rose (1990), Krautmann (1990), Hill (1985), Durland and Sommers (1991), Sommers (1993), Krautmann and Oppenheimer (1994), Krautmann (1999), Krautmann, Gustafson, and Hadley (2000), Maxcy, Fort, and Krautmann (2002), Krautmann and Oppenheimer (2002), and Goff, McCormick, and Tollison (2002). Sommers (1993) also used a player's batting average, whereas Krautmann, Gustafson, and Hadley (2000) added a hitter's runs-batted-in. Slugging percentage has not been the only measure of productivity chosen. Medoff (1976), Hill and Spellman (1983), and MacDonald and Reynolds (1994) measured a hitter's productivity with runs scored.
3. The data used to estimate equation 1 was taken from baseball-reference.com. The data set includes all Major League Baseball teams from 1996 to 2008. Results available upon request. p. 14, note #6: Replace last sentence with "Previous research confirms that OPS is superior to slugging average for measuring hitter performance.

4. Slugging average data also come from baseball-reference.com and our results for slugging also are available on request.
5. See Krautmann, Von Allmen, and Berri (2009) for a recent example.
6. Andrew Zimbalist (1992b) also argued that "PROD" (which is OPS) was superior to slugging percentage. Later research confirmed that Zimbalist approach with respect to hitters was better.
7. Although McCracken was not aware of this, James (1987) had also sought to evaluate pitchers with a DIPS-like metric called "indicated ERA," which judged pitchers by walks and home runs allowed. However, James did not make the insight that McCracken did regarding pitchers having little control over balls in play.
8. Bradbury (2007b) finds that the pitchers' market was properly valuing pitching talent according to DIPS metrics even before McCracken's findings were published.
9. Heeran (1992) begins with a model identical to the one currently used by the NBA but then weights each player's production by both minutes played and the average game pace his team played throughout the season being examined. More information on NBA Efficiency can be found at NBA.com [Efficiency: The Daily Recap" http://www.nba.com/statistics/efficiency.html].
10. Bellotti (1996) begins with the basic TENDEX model and then simply subtracts 50% of each player's personal fouls. Jeffrey Jenkins (1996) used the Points Created model in a study of racial discrimination.
11. sports.espn.go.com/nba/columns/story?columnist=hollinger_john&id=2850240
12. A player who makes 2.9 shots from two-point range would get 5.8 points. Plus a player gets 0.4 for each field goal made. So 2.9 made shots increases Game Score by 1.2. Adding 5.8 and 1.2 together, we see that the player who shoots 29% from two-point range just breaks even.
13. An alternative testing approach was suggested and originally presented to the online Apbermetric community by Lewin and Rosenbaum (2007). Their approach began with a team's efficiency differential or points scored per possession minus points surrendered per possession. This differential was then regressed on a team's PERs (or whatever metric these authors were examining). Then, using the results of this regression, plus the residual from the regression, each player was evaluated. This evaluation was then used to predict the next season's efficiency differential for each team (with a rookies' performance simply taken as given). The results indicated that all the models examined were able to explain between 75% and 77% of future wins. In other words, all models were basically the same. Such a result should not be surprising. As Lewin and Rosenbaum (2007) actually argue, any model plus the residual does an equally good job of explaining current wins. Of course, such an argument applies to all models (as any student of econometrics should understand). In sum, it is not clear what the Lewin-Rosenbaum residual approach is designed to tell us. It is certainly well understood that any model plus the residual can explain 100% of a dependent variable. And as we see, when we do not use the residual in our evaluation, we do see clear differences in the explanatory power of different metrics.
14. The Wins Produced model was also discussed by Berri, Schmidt, and Brook (2006). It builds upon what was presented by Berri and Krautmann (2006) and Berri (1999).

15. One can go one step further and insert the individual variables that comprise the team defensive adjustment into equation (3) separately. Such a step does raise explanatory power for NBA Efficiency, PERs, and Game Score to 82%, 82%, and 85%. Remember, Wins Produced explains 94% of wins. So even with the team defensive adjustments components added separately, the "efficiency" metrics come up short.
16. This definition is taken from Hockey-Reference.com.
17. Our result comes from a regression that was ran on 2,729 skater observations taken from the 2001-02 to 2007-08 season. Skaters had to have a minimum of 500 min of ice time in both the current and lagged season to be included in the sample. Data were taken from Hockey-Reference.com. Results available on request.
18. The explanatory power of each statistic is as follows: shooting percentage, 39%; assists, 55%; goals, 63%; points, 69%; penalty minutes, 71%; and shots on goal, 80%. The data set included 2,729 skater observations taken from the 2001-02 to 2007-08 season. Skaters had to have a minimum of 500 minutes of ice time in both the current and lagged season to be included in the sample. Data were taken from Hockey-Reference.com. Results available on request.
19. For basketball, 82games.com reports a player's Net48, which is defined as "the team net points per 48 min of playing time for the player." "Net points" is the plus/minus statistic for basketball or the difference between the points a team scores and allows when a player is on the court. The consistency of Net48 was established with data on 364 players from the 2006-07 to 2008-09 seasons. The player had to play at least 1,000 min in consecutive seasons to be included in the data set. 82games.com also reports Net On Court/Off Court. A player's off-court performance is the net points a team realizes when the player is not in the game. On court is simply Net48. So Net on court/off court is intended to capture how well a team performs with and without the player. Only 12% of a player's Net on court/off court is explained by what the player did last year. Results available on request.
20. The data we are using for our discussion on adjusted plus–minus come from BasketballValue.com. These data were compiled by Aaron Barzilai. According to the BasketballValue.com, the calculations were done in the spirit of the work of Dan Rosenbaum. Rosenbaum's work, in turn, is based on the work of Wayne Winston and Jeff Sagarin. We do not have access to the original work of Winston-Sagarin; so, we cannot say the issues we raise apply to the work of Winston-Sagarin.
21. Given the issues with adjusted plus–minus, it is odd that Lewin and Rosenbaum (2007) sought to "test" various models (i.e., PERS, Wins Produced, etc. . . . .) by looking at the correlation between player evaluations completed with adjusted plus–minus evaluations and how players were evaluated by other metrics. Such a "test" appears to ignore the problems we cite with adjusted plus–minus.
22. Published research in this field includes the work of Berri and Krautmann (2006), Berri, Schmidt, and Brook (2006), and Berri (2008). This work details the Wins Produced model. This model, based on box score statistics, explains 94% of current wins and is quite consistent over time. Wins Produced per 48 min has a .75 correlation across time (based on data from the 1977-78 to 2007-08 seasons). If one removes the position adjustment from Wins Produced, the correlation rises to .83.

Declaration of Conflicting Interests
The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding
The authors received no financial support for the research and/or authorship of this article.

Web Sites

82games.com: created by Roland Beech. The site contains plus–minus analysis of NBA players.

Baseball Prospectus (baseballprospectus.com): created by Gary Huckaby and is run by Prospectus Entertainment Ventures, LLC. The site contains commentary and proprietary statistics for evaluating baseball performance.

Baseball-Reference (baseball-reference.com): created and operated by former mathematics professor Sean Forman. This site contains historical Major League and Minor League Baseball statistics.

BasketballValue.com: created by Aaron Barzilai. The site contains adjusted plus–minus analysis of NBA players.

Basketball-Reference.com (basketball-reference.com): created by Justin Kubatko. The site provides statistical data on every player and team in NBA, American Basketball Association (ABA), and Women's National Basketball Association (WNBA) history.

Hawkonomics (hawkonomics.blogspot.com): created by Stacey Brook. The site provides commentary on sports and other economic issues.

Pro-Football-Reference.com (pro-football-reference.com): created by math professor Doug Drinen. This site contains historical professional football statistics.

Sabernomics (sabernomics.com): created by economics professor (and coauthor) J. C. Bradbury. The site contains commentary relating to the economics of baseball and sabermetrics.

The Sports Economists (thesportseconomists.com): created by economics professor Raymond Sauer. The site contains commentary on sports economics by many economics professors.

The Wages of Wins Journal (dberri.wordpress.com): created by economics professor (and coauthor) David Berri. The site contains stories that follow the theme of The Wages of Wins.

References

Albert, J., & Bennett, J. (2003). *Curve ball: Baseball, statistics, and the role of chance in the game.* New York: Copernicus.

Bellotti, R. (1996). *The points created basketball book, 1992-93.* Night Work Publishing Co.: New Brunswick.

Berri, D. J. (1999, Fall). Who is most valuable? Measuring the player's production of wins in the National Basketball Association. *Managerial and Decision Economics*, *20*, 411-427.

Berri, D. J. (2008). A simple measure of worker productivity in the National Basketball Association. In B. Humphreys, & D. Howard (Eds.), *The Business of Sport* (3 vol, pp. 1-40). Westport, CT: Praeger.

Berri, D. J., & Krautmann, A. (2006, July). Shirking on the court: Testing for the dis-incentive effects of guaranteed pay. *Economic Inquiry*, *44*, 536-546.

Berri, D. J., Schmidt, M. B., & Brook, S. L. (2006). *The wages of wins: Taking measure of the many myths in modern sports.* Stanford, California: Stanford University Press.

Birnbaum, P. (2006). "Chopped Liver," Sabermetric Research Web site. Retrieved on November 11, 2009, from http://sabermetricresearch.blogspot.com/2006/10/chopped-liver.html

Blass, A. A. (1992). Does the Baseball Labor Market contradict the human capital model of investment? *The Review of Economics and Statistics*, *74*, 261-268.

Bradbury, J. C. (2007a). *The Baseball economist: The real game exposed*. Plum, New York: Dutton.

Bradbury, J. C. (2007b). Does the Baseball Labor Market properly value pitchers? *Journal of Sports Economics*, *8*, 616-632.

Bradbury, J. C. (2008). Statistics performance analysis in sport. In B. R. Humphreys, & D. R. Howard (Eds.), *The business of sport: Volume 3: Bridging research and practice* (pp. 41-56). Westport, CT: Praeger.

Bruggink, T. H., & Rose, D. R. Jr. (1990, April). Financial restraint in the free agent labor market for Major League Baseball: Players look at strike three. *Southern Economic Journal*, *56*, 1029-1043.

Durland, D. Jr., & Sommers, P. M. (1991, March). Collusion in Major League Baseball: An empirical test. *Journal of Sport Behavior*, *14*, 19-29.

Goff, B. L., McCormick, R. E., & Tollison, R. D. (2002, March). Racial Integration as an innovation: Empirical evidence from sports leagues. *American Economic Review*, *92*, 16-26.

Heeren, D. (1992). *Basketball abstract, 1991-92 edition.* Englewood Cliffs, NJ: Prentice Hall, Inc.

Hill, J. R. (1985, Winter). The threat of free agency and exploitation in professional baseball: 1976-1979. *Quarterly Review of Economics and Business*, *25*, 68-82.

Hill, J. R., & Spellman, W. (1983, Winter). Professional baseball: The reserve clause and salary structure. *Industrial Relations*, *22*, 1-19.

Hollinger, J. (2002). *Pro basketball prospectus 2002.* Washington, DC: Brassey's Sports.

James, B. (1987). *The Bill James Baseball Abstract, 1987.* New York: Ballantine Books.

Jenkins, Jeffery A. 1996. A Reexamination of Salary Discrimination in Professional Basketball. *Social Science Quarterly 77*, no. 3 (September): 594-608.

Jewell, R. T. (2006). Sports economics: State of the discipline. In J. Fizel (Ed.), *Handbook of Sports Economics Research.* Armonk, NY: M.E. Sharpe. pp. 9-18.

Kahn, L. M. (2000, Summer). The sports business as a labor market laboratory. *Journal of Economic Perspectives*, *14*, 75-94.

Krautmann, A. (1990, April). Shirking or stochastic productivity in Major League Baseball? *Southern Economic Review*, 961-968.

Krautmann, A. (1999, April). What's wrong with Scully-estimates of a player's marginal revenue product. *Economic Inquiry*, *37*, 369-381.

Krautmann, A., Gustafson, E., & Hadley, L. (2000, January). Who pays for minor league training costs? *Contemporary Economic Policy*, *18*, 37-47.

Krautmann, A., & Oppenheimer, M. (2002, February). Contract length and the return to performance in Major League Baseball. *Journal of Sports Economics*, *3*, 6-17.

Krautmann, A., & Oppenheimer, M. (1994, September–October). Free agency and the allocation of labor in Major League Baseball. *Managerial and Decision Economics*, *15*, 459-469.

Krautmann, A., Von Allmen, P., & Berri, D. J. (2009). The underpayment of restricted players in North American Sports Leagues. *International Journal of Sport Finance*, *4*, 155-169.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*, 1121-1134.

Lewin, D., & Rosenbaum, D. (2007). *The pot calling the kettle black: Are NBA statistical models more irrational than "irrational" decision-makers?* Unpublished manuscript.

Lindsey, G. (1963). An investigation of strategies in baseball. *Operations Research*, *7*, 477-501.

MacDonald, D. N., & Reynolds, M. O. (1994, September–October). Are baseball players paid their marginal product? *Managerial and Decision Economics*, *15*, 443-457.

Maxcy, J., Fort, R., & Krautmann, A. (2002, August). The effectiveness of incentive mechanisms in Major League Baseball. *Journal of Sports Economics*, *3*, 246-255.

McCracken, V. (2001). Pitching and defense: How much control do hurlers have? Baseball Prospectus (Web site). Retrieved on November 11, 2009 from http://baseballprospectus.com/article.php?articleid=878.

Medoff, M. H. (1976). On monopsonistic exploitation in professional baseball. *Quarterly Review of Economics and Business*, *16*, 113-121.

NBA.com. Efficiency: The daily recap. Retrieved from http://www.nba.com/statistics/efficiency.html

Oliver, D. (2004). *Basketball on paper.* Washington DC: Brassey's Inc.

Peach, J. (2004, June). Thorstein Veblen, Ty Cobb, and the evolution of an institution. *Journal of Economic Issues*, *XXXVIII*, 327-338.

Porter, P. K., & Scully, G. W. (1982). Measuring managerial efficiency: The case of baseball. *The Southern Economic Journal*, *48*, 642-650.

Raimondo, H. J. (1983, Spring). Free agents' impact on the labor market for baseball players. *Journal of Labor Research*, *4*, 183-193.

Scully, G. W. *The Business of Major League Baseball*. Chicago: The University of Chicago Press, 1989.

Scully, G. W. (1974). Pay and performance in Major League Baseball. *American Economic Review*, *64*, 917-930.

Scott, F. Jr., Long, J., & Sompii, K. (1985). Salary vs. marginal revenue product under monopsony and competition: The case of professional basketball. *Atlantic Economic Journal*, *13*, 50-59.

Sommers, P. M. (1990, December). An empirical note on salaries in Major League Baseball. *Social Science Quarterly*, *71*, 861-867.

Sommers, P. M. (1993, June). The influence of salary arbitration on player performance. *Social Science Quarterly*, *74*, 439-443.

Sommers, P. M., & Quinton, N. (1982, Summer). Pay and performance in Major League Baseball: The case of the first family of free agents. *Journal of Human Resources*, *17*, 426-436.

Thorn, G., & Palmer, P. (1984). *The hidden game of baseball: A revolutionary approach to baseball and statistics.* New York: Dolphin.

Turocy, T. L. (2005, December). Offensive performance, omitted variables, and the value of speed in baseball. *Economics Letters*, *89*, 283-286.

Veblen, T. 1899. *The Theory of the Leisure Class*. Reprint, New York: The New American Library, 1953.

Zimbalist, A. (1992a). Salaries and performance: Beyond the Scully model. In P. Sommers (Ed.), *Diamonds are forever: The business of baseball* (pp. 109-133). The Brookings Institution: Washington, DC.

Zimbalist, A. (1992b). *Baseball and billions.* Basic Books: New York.

Bios

David J. Berri is an associate professor of economics in the Department of Economics and Finance at Southern Utah University. His current research focuses on the economics of sport, specifically the topics of consumer demand, competitive balance, and worker productivity.

John Charles Bradbury is an associate professor of health, physical education, and sport science at Kennesaw State University. He received his PhD in economics from George Mason University in 2000. His research interests include public economics and the economics of sports. He is the author of *The Baseball Economist* (Dutton 2007).