

The Pot Calling the Kettle Black

Are NBA Statistical Models More Irrational than “Irrational” Decision-Makers?

September 2007

David Lewin (Macalester College)

and

Dan T. Rosenbaum (UNC-Greensboro)

Acknowledgments: We would like to thank the APBRmetrics community for years of commentary and critique of the ideas contained in this paper. We also would like to make it clear that this paper represents the views of the authors alone and not the institutions that the authors are associated with.

Stats Better than Human Judgment

- Michael Lewis (*Moneyball*)
- Malcolm Gladwell (review of *Wages of Wins*)
- David Leonhardt (*New York Times*)

“Academic research, however, is pretty much on the side of statistics. Whether diagnosing patients or evaluating job candidates, human beings vastly overestimate their ability to make judgments, research shows. Numbers and analysis almost always make people better.

‘There have been hundreds of papers on subjects from picking students for a school to predicting the survival of cancer patients,’ said Richard Thaler, a University of Chicago economist who uses sports examples in his class on decision-making. When a computer model is given the same information as an expert, the model almost always comes out on top, Thaler said.”

(Leonhardt, 2005).

Are NBA Decision-Makers Rational?

- David Berri et al. (*Wages of Wins*, 2006, p. 199)
“Our story of the overrated and underrated indicates that the NBA may have a problem evaluating talent. The overrated players can all score, and most of these players have also scored major paydays. . . All of this suggests that people making decisions in the NBA are not as “rational” as economists tend to expect.”
- Dan Rosenbaum (*New York Times*, 2005)
“Teams pay for little more than the glory statistics (points, rebounds and, to a lesser extent, assists). . . Although steals, blocks, shooting percentage and an ability to avoid turnovers are crucial to a team’s performance, players proficient in these aspects are rarely rewarded with bigger paychecks.”

NBA Production Functions

- “When a computer model is given the *same information* as an expert, the model almost always comes out on top, Thaler said.” (Leonhardt, 2005).
- Does the computer model ever have the “same information” as the expert?
 - In baseball that may be close to true.
 - In basketball probably not, given the difficulties of apportioning credit (although this is a problem for the expert too) and the lack of data on things like setting a good pick, playing help defense, and spreading the court.
 - Advanced statistical models often do much worse than the naïve models of NBA decision-making.

NBA Production Functions Cont.

- How do we model NBA decision-making, i.e. what is the production function they have in mind?
 - Minutes per Game
 - Points per Game
 - NBA Efficiency (add up good things and subtracts bad things)
 - Add in team adjustment
- What are the alternative production functions that statistical analysts assume?
 - Wins Produced: assumes team production function is applicable to individual players
 - PER: a more careful and reasoned weighting of good and bad things
 - Adjusted Plus/Minus: a non-box score metric

Adjusted Plus/Minus

- Measures how the point differential changes when a given player is in the game, holding the effects of the other players (on both teams) constant.
- In theory this should pick up almost all of the contributions a player makes, including picks, good defense, spreading the floor, etc.
- Not widely used because of difficulty of computing.
 - Games are broken down into all of the combinations of players from both teams (over 182,000 combinations between 2002-03 and 2006-07).
 - Can't be easily done in Excel and requires solid understanding of regression.

Adjusted Plus/Minus Continued

$\text{DIFF} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K + \varepsilon$, where

- DIFF = home team minus away team points per possession.
- $X_i = 1$ if player i is playing at home, $= -1$ if player i is playing away, $= 0$ otherwise (i.e. if player i is not playing),
- ε = i.i.d. error term,
- β_0 measures the average home court advantage across all teams,
- $\beta_1, \beta_2, \dots, \beta_K$ measure how the *team* point differential changes when a particular player is on the court (relative to the reference players), holding the effects of other players (on both teams) constant.
 - The reference players are all players playing less than 250 minutes in a given season.

Table 1

Correlations of Player Metrics with Box Score Statistics Correlation Coefficient (Standard Error)

Player Metric	Per 40 Minute Statistics							Team EFF
	PTS	TSA	REB	AST	TO	STL	BLK	
Minutes per Game (MPG) with team adjustment	0.466 (0.029)	0.403 (0.031)	0.122 (0.044)	0.352 (0.038)	0.111 (0.048)	0.048 (0.033)	0.096 (0.044)	0.654 (0.011)
Points per Game (PPG) with team adjustment	0.846 (0.019)	0.801 (0.025)	0.099 (0.057)	0.402 (0.068)	0.401 (0.054)	0.048 (0.051)	0.069 (0.051)	0.311 (0.009)
NBA Efficiency (EFF) with position and team adjustments	0.724 (0.045)	0.631 (0.050)	0.488 (0.059)	0.508 (0.069)	0.368 (0.060)	0.216 (0.050)	0.294 (0.052)	0.329 (0.007)
Player Efficiency Rating (PER) with position and team adjustments	0.833 (0.035)	0.744 (0.043)	0.327 (0.063)	0.483 (0.071)	0.382 (0.062)	0.187 (0.049)	0.237 (0.054)	0.242 (0.009)
Wins Produced (WP) with position and team adjustments	0.301 (0.062)	0.181 (0.063)	0.684 (0.045)	0.401 (0.062)	0.050 (0.062)	0.287 (0.048)	0.352 (0.055)	0.341 (0.009)
Alt. Wins Produced (AWP) with position and team adjustments	0.590 (0.048)	0.468 (0.053)	0.419 (0.058)	0.446 (0.063)	0.116 (0.063)	0.252 (0.043)	0.270 (0.056)	0.373 (0.008)
Adjusted Plus/Minus (+/-)	0.400 (0.039)	0.349 (0.043)	0.126 (0.051)	0.335 (0.046)	0.166 (0.037)	0.162 (0.037)	0.112 (0.054)	0.115 (0.017)

- MPG, PPG, EFF, and PER all give scoring a lot of weight. WP give a lot more weight on rebounds than other metrics. Correlation coefficients for +/- similar to those for PPG. Note that per 40 minute statistics are position-adjusted.

Putting Stats to the Test

- Explaining current team wins using current metrics.
 - With team adjustments, this tells us nothing.
- Predicting current team wins using past metrics.
 - Assume that metrics give per minute productivity.
 - Assume minutes played can be perfectly predicted.
 - Assume perfect prediction for low minutes players/rookies.
 - Predicting wins from 1980-81 through 2006-07 seasons.
- Predicting current adjusted plus/minus using past metrics.
 - More powerful test than using team wins.
 - Predicting +/- from 2002-03 through 2006-07 seasons.

Table 2

Correlations of Player Metrics with Team Wins and +/- Correlation Coefficient (Standard Error)

Player Metric	Adjustments		Predicting	
	Pos.	Team	Wins	+/-
Player Efficiency Rating (PER)	Yes	Yes	0.805 (0.022)	0.492 (0.043)
Alt. Wins Produced (AWP)	Yes	Yes	<i>0.826</i> <i>(0.020)</i>	0.457 (0.040)
NBA Efficiency (EFF)	Yes	Yes	0.818 (0.021)	0.479 (0.040)
Points per Game (PPG)	No	Yes	0.818 (0.020)	0.456 (0.039)
Minutes per Game (MPG)	No	Yes	0.823 (0.017)	0.351 (0.023)
Wins Produced (WP)	Yes	Yes	0.801 (0.023)	0.372 (0.048)
Adjusted Plus/Minus (+/-)	No	No	0.102 (0.016)	0.267 (0.033)

Correlation coefficients for player metrics that are significantly different from Wins Produced at the 5% (1%) level are in italics (bold).

Conclusion

- Berri et al. 2006, p. 215
“One can play basketball. One can watch basketball. One can both play and watch basketball for a thousand years. If you do not systematically track what the players do, and then uncover the statistical relationship between these actions and wins, you will never know why teams win and why teams lose. Staring at these players play is not a method that will ever yield the answers that the proper analysis of statistics will yield. And this is true if stare for one day, or as we said, if you stare for a thousand years.”
- One lesson is that we should have some evidence before making a claim like this one from Berri et al.
- But the bigger lesson is that NBA statistical analysts can make contributions, but only if we realize that our models can be worse than “staring.”

Table 3

Top 10 Lists for 2005-06

	MPG	PPG	PER	EFF
1	Dirk Nowitzki	Kobe Bryant	Kobe Bryant	Dwyane Wade
2	Tim Duncan	LeBron James	Dwyane Wade	Kevin Garnett
3	Tony Parker	Allen Iverson	LeBron James	Kobe Bryant
4	Bruce Bowen	Dirk Nowitzki	Dirk Nowitzki	LeBron James
5	Jason Terry	Dwyane Wade	Yao Ming	Yao Ming
6	Chauncey Billups	Gilbert Arenas	Kevin Garnett	Dirk Nowitzki
7	Pau Gasol	Carmelo Anthony	Elton Brand	Elton Brand
8	Dwyane Wade	Paul Pierce	Shaquille O'Neal	Tim Duncan
9	Shawn Marion	Vince Carter	Allen Iverson	Shawn Marion
10	LeBron James	Michael Redd	Manu Ginobili	Shaquille O'Neal

	WP	AWP	+/-
1	Kevin Garnett	Kevin Garnett	Ray Allen
2	Jason Kidd	Dirk Nowitzki	Kobe Bryant
3	Shawn Marion	Dwyane Wade	Andre Miller
4	Marcus Camby	Elton Brand	Ben Wallace
5	Ben Wallace	LeBron James	LeBron James
6	Dwyane Wade	Kobe Bryant	Jason Richardson
7	Tim Duncan	Shawn Marion	Mehmet Okur
8	Jeff Foster	Manu Ginobili	Dwight Howard
9	Manu Ginobili	Jason Kidd	Rasheed Wallace
10	Steve Nash	Steve Nash	Vince Carter