THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF STATISTICS

DEVELOPING A NEW METRIC FOR NBA PLAYER EVALUATION: POINTS PER LOST POSSESSION

DEREK CHARLES GERBERICH
SPRING 2014

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree in Statistics
with honors in Applied Statistics

Reviewed and approved* by the following:

Andrew Wiesner
Lecturer of Statistics
Thesis Supervisor

Murali Haran
Associate Professor of Statistics
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

The basic statistical box-score measures of the NBA have been recorded since the game's inception in the late 1800's. Is it not intriguing that we still evaluate player performance by the same measures that were used nearly a century ago? Surely there must be room for improvement. Where is the Bill James of the NBA? What is the NBA equivalent to the MLB's On Base Percentage? Finally, how would one go about developing such a metric? Presumably, a new statistical insight into one of American's most popular and lucrative ($4.3 billion) industries would be quite valuable. Can we improve upon a measure that has been accepted for more than 100 years? This paper aims to explore that possibility.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

-Psalm 106:1
"Praise the Lord. Give thanks to the Lord, for He is good; His love endures forever."

-To my dad, who gave me a basketball just eighteen months into my life, and introduced me to an excel spreadsheet at age four.

-To my mom, for seeking the Lord through prayer for His wisdom and counsel in raising her children.

-To Dr. Wiesner, the world needs more people like you.

-To Ricky Mouser—I will always cherish our late night (early morning?) talks about combinatorics and basketball theory.

Introduction

Moneyball and Professional Sports

In 2003, the book *Moneyball: The Art of Winning an Unfair Game* rocked the sports analytics landscape, as it unpacked the advanced statistical thinking behind the Oakland Athletics and their unlikely success as a small-market team in Major League Baseball. In a profession which is dominated by big-market teams simply exercising their will by outspending all of their counterparts, the small-market and small budget Athletics managed to pull off improbable success, making it to the American League Division Championship in four consecutive seasons from 2000-2003. For most baseball observers, this unprecedented success from a small-market team left them scratching their heads. How could an afterthought of a team, that didn't have the resources or payroll to compete with the most powerful teams in the league, find itself surrounded by consistent success? The answer, as we would find out, was hidden in the numbers (Surowiecki).

Bill James

Bill James, considered "one of the most prominent authors on the subject of baseball from a statistical perspective" invented what is now termed today as sabermetrics, with his "insightful, creative, and comical writings on a plethora of baseball topics" (Bill James). Now employed as the Senior Baseball Operations Advisor of the Boston Red Sox, James revolutionized the Major League Baseball industry with his unique approach to analyzing the numbers that had been recorded and used in analysis for nearly a century.

A Case Study: The Oakland Athletics

To illustrate the simple yet profound contributions of Bill James, let's go back to the Oakland Athletics example. One of the keys to their success was evaluating players based on simple metric tweaks. By applying these new metrics, the Athletics were able to identify and exploit market inefficiencies.

One of the metrics that the Athletics employed was examining a player's On Base Percentage (OBP). At that point, what had traditionally been the norm for teams to measure a player's batting prowess by was simply his Batting Average (BA). A batting average records the percentage of time that a player gets a hit when he is at bat. However, the simple brilliance of OBP was that it recorded the percentage of time a player got *on base* per at bat, (rather than just getting a hit). In baseball, getting on base provides the same value whether you get a hit, walk, or are hit by a pitch. If you end up on base, you get on base. Therefore, players who were apt at drawing walks were undervalued in the Major League Baseball market at that time. The Oakland A's evaluation of a player by his OBP instead of his BA, allowed them to spot player bargains, and get the most production out of their constrained payroll. Such measures eventually allowed the small-market Athletics to prosper, and in the end reap the rewards of success in a $7.2 billion industry (The Business of Sports).

Further Applications

The remarkable part of the story is that the success achieved by statistical insights has yet to be parlayed to any other of the major sports. Let's take for example the National Basketball Association (NBA). The basic statistical box-score measures of the NBA have been recorded since the game's inception in the late 1800's. Is it not intriguing that we still evaluate player performance by the same measures that were used nearly a century ago? Surely there must be room for improvement. Where is the Bill James of the NBA? What is the NBA equivalent to the MLB's On Base Percentage? Finally, how would one go about developing such a metric? Presumably, a new statistical insight into one of American's most popular and lucrative ($4.3 billion) industries would be quite valuable. Can we improve upon a measure that has been accepted for more than 100 years? This paper aims to explore that possibility.

Chapter 1: Literature Review of Existing Methods

How Are NBA Players Paid?

In 2007, a thought provoking article was published in the International Journal of Sports Finance entitled: “Does One Simply Need to Score to Score?” which evaluated whether or not an individual player could earn a big paycheck if all he focused on was scoring the basketball. In summation, the final answer to the question was a resounding yes. If an NBA player was simply to focus on only scoring the basketball, he would always end up well paid and employed in the NBA (Berri, Brook and Schmidt). Why is this true?

The reality is, that despite all of the statistical data available to the NBA, and NBA personnel decision makers, players continually end up earning money and playing time simply by the amount of points they score per game. It is a flaw which does not allow NBA teams to maximize their potential to win games. Because of this flaw, the authors of the paper argue that “decision makers in the NBA do not behave according to the dictates of instrumental reality.” In summation, NBA decision makers in the front office do not behave rationally when it comes to building their teams because they use points scored as a leading metric to evaluate their players. In an industry that is driven by a success/failure bottom line, not operating at one’s maximum level of efficiency can cause massive losses in employment, as well as significant monetary casualties.

Consequential Questions

This finding leads to two questions:

(1) What makes points per game such a poor measure, and can we show that it truly is poor?

(2) If points per game are truly a poor statistical measure, why do NBA front offices use it as a driving factor in personnel acquisition?

Let's quickly address the answer to question number two (2) first, and then move into a more lengthy discussion over number one (1). If points per game are truly a poor player evaluator, why is it used across NBA front offices? The answer is that NBA teams do not have a better alternative. There has been no Bill James to revolutionize the NBA. Front offices across the league are stuck using a metric that is as old as the game itself, and paying their prized commodities (the players) heavily based upon how many points they produce. It's a system that is open for improvement.

Of course there are some who would argue that points per game are an effective evaluation metric. The fact is that is has been used for over 100 years. Which leads us back to our first question (1)—what makes points per game such a poor metric, and can we statistically show its inadequacies?

The beginnings of that answer come from basketball intuition. Yes, scoring points is an important measure, and yes, you cannot win a basketball game without scoring. But the overlooked fact is that every point scored comes with a price—a shot attempt. In general, one must shoot to score. And when one shoots, not all shots end in made baskets. But theoretically, they could. And theoretically every missed shot comes with the opportunity cost of another (most likely better) shot, that could have been available had the shooter passed up on the shot that caused the missed attempt. Additionally, in the pursuit of scoring points, players make errors, and turn the ball over. Again in this situation, turning the ball over also comes with a direct cost. When the ball is turned over, a possession is lost in pursuit of scoring points. To summarize all of this, all points scored are not created equal. For a simple illustration, let's take a look at a sample game of two players below:

Player A: 16 PTS, (8 shots made on 20 attempts), 4 turnovers

Player B: 16 PTS, (8 shots made on 12 attempts), 1 turnover

It's fairly obvious from this comparison, that you would rather have Player B on your team. Player B managed to score the same amount of points as Player A, while shooting the ball eight less times, and committing three less turnovers. Yet, when an NBA general manager is evaluating the players based only on points scored, both would produce an average of 16 points per game, and both would be paid smartly for their services. But shouldn't Player B be considered the more valuable commodity? Well, yes. But how can we determine a rigorous and reproducible metric to address this issue?

Points Per Miss (PPM)

A line of thinking into the issue begins with an interesting column posted by Jerod Morris, a sports writer for Midwest Sports Fans (Morris). Morris's article begins to dig at the core of the matter—not all points are created equal, and points come at a price. He tosses out a metric from Bill Simmons which he calls Points Per Miss (PPM). Here, looking at how many points a player scores for every missed shot they incur, Morris tries to evaluate a player beyond simply points scored. It's an idea that is on the right track, but what can we do about turnovers? Similarly, are missed free throws the same thing as missed field goals? Lastly, how can we factor offensive rebounds (when the team who misses the shot gets the rebound) into a metric? An adequate discussion and potential solution is presented in the following chapters.

Chapter 2: Development of Points Per Lost Possession (PPLP)

A Proposition

To answer the questions and challenges brought up so far in this paper, let us propose a new NBA player evaluation metric: Points Per Lost Possession (PPLP). To define the metric, we'll start with constructing a template for what we intend to produce. When constructing the metric, we want our final output to take the shape of a fraction defined as:

{Points} / {Lost Possessions}

To define points is simple. For our numerator, we will include the total amount of points scored by an individual (or team) over the given time period of evaluation, for example a full NBA season. What becomes the challenging aspect is to define a lost possession.

Lost Possession

For the purposes of this paper, let us define a lost possession as being contrived from the summation of three basic parts:

(1) A turnover

(2) A missed field goal that is rebounded by the opposing team

(3) A missed free throw that is rebounded by the opposing team

These three parts would combine to total the number of lost possessions, which can be calculated for both a team and an individual player. It is worth noting that theoretically speaking, a player or team could have zero lost possessions (if they made every shot, and never turned the ball over). In this case, our ratio would be left undefined. However, while this is theoretically possible, it is extremely unlikely to occur.

Turnover

A turnover occurs when the offensive team (the team with the ball) commits an act which causes the ball to be given (turned over) to the opposing team. Turnovers are tracked in every NBA game, and can be summed for both individual players and teams.

Missed Field Goal (Rebounded by the Opposing Team)

A missed field goal is a shot attempt from the playing field which does not go into the basket. To determine if the shot was rebounded by the opposing team, we will take the amount of missed field goals by a team and multiply that number by the complement of the team's respective offensive rebounding rate. When calculated for individual players, the offensive rebound rate used in the calculations will be the league average over the past decade (~27%) (NBA Offensive Rebounding). The formula is as follows:

Missed Field Goal (Rebounded by Opposing Team)

= Field Goals Missed * [1 – Offensive Rebound %]

For example, if your team rebounds 30% of its missed shots, the number of lost possessions per 100 missed shots would be:

= 100 * (1 - .30) = 70 lost possessions

Missed Free Throw (Rebounded by the Opposing Team)

The third part of our summation lies in defining a missed free throw that is rebounded by the opposing team. Defining this is the most complex part of the process. While we can total the number of free throws that are missed by a team or individual, all missed free throws are not the same. Some missed free throws come with an opportunity for an offensive rebound, while others do not. The breakdown of free throws with the possibility to be rebounded versus free throws with no possibility of being rebounded is as follows[1]:

Possibility of Being Rebounded

- 2nd shot of a 2-shot foul
- 3rd shot of a 3-shot foul
- All 1-shot fouls (In basketball terms an "and-1").

No Possibility of Being Rebounded

- 1st shot of a 2-shot foul
- 1st shot of a 3-shot foul
- 2nd shot of a 3-shot foul

In addition to trying to account for the variety of free throws that provide the possibility for an offensive rebound, one also must consider that the offensive rebound rate for a free throw is different than the offensive rebound rate for a regular field goal attempt. The reasoning behind this is because in a regular field goal attempt, the alignment of players is more scattered, and any player can break for the rebound at any point. A free throw, however, is more structured. When set up, there are four defensive

[1]Note: there are no one-and-one foul shots in the NBA.

players allowed to be in the lane for the rebound compared to two offensive players. The players may also not break for the rebound until the shot is released. This creates a more structured and controlled environment for rebounding, making it harder to obtain an offensive rebound from a missed free throw compared to a missed field goal. The question remains how much harder is it to obtain an offensive rebound via a missed free throw compared to a missed field goal?

Thanks to statistics tracked by 82games.com, we know that the offensive rebound rate for missed free throws is 14%, which is significantly less than a missed field goal (NBA Offensive Rebuilding Studies). But when it comes to the breakdown of which free throw attempts can be possibly rebounded, the research becomes a bit more complex. From TeamRankings.com we see that the average amount of fouls per NBA game is approximately 20 per game (NBA Basketball). Using 82games.com foul tracking output, we see that NBA fouls are distributed as such:

44% of fouls occur while shooting a 2-pointer

2% of fouls occur while shooting a 3-pointer

50% of fouls occur in the backcourt

4% of fouls occur in an “and-1” scenario

With this distribution of fouls, we encounter another issue. How many of the fouls that occur in the backcourt result in free throws? Again, from 82games.com, we find that backcourt fouls end up contributing 11.72 foul shots per game, to the overall total. All of those shots are two-shot fouls. With that in mind, we can construct the following distribution which would break down the percentage of free throws that are available to be rebounded in any given NBA game (NBA Game Charting):

- Fouled while shooting a 2-point field goal:

  [44%] = 8.8 fouls per game

  = 17.6 shots per game

  [8.8 shots have a chance to be rebounded, 8.8 shots do not]

- Fouled while shooting a 3-point field goal:

  [2%] = 0.4 fouls per game

  = 1.2 shots per game

  [0.8 shots have a chance to be rebounded, 0.4 shots do not]

- Fouled in the backcourt:

  [50%] = 11.72 shots per game

  [5.9 shots have a chance to be rebounded, 5.9 shots do not]

- Fouled while making a field goal ("and-1"):

  [4%] = 0.8 fouls per game

  = 0.8 shots per game

  [0.8 shots have a chance to be rebounded, 0.0 shots do not]

When we total the above numbers, we end up with an average of 15.5 free throws per game that have no possibility of being rebounded, and 15.9 free throws per game that have the possibility to be rebounded (if they are missed). That means that approximately 15.9/31.4 = 51% of all free throws attempted will present an opportunity for an offensive rebound if they are missed. For our formula, we will assume that the rate of missed free throws is the same whether or not the shot is going to be available for a potential rebound or not.

After this research and calculations, we can now define our lost possession for a missed free throw (rebounded by the opposing team):

Missed Free Throw (Rebounded by Opposing Team)

= [Free Throws Missed * 0.5 * 0.49] +

[Free Throws Missed * 0.5 * (1-0.14) * 0.51]

To break this down, we see that 49% of all missed free throws have no chance of being rebounded, and a half of possession consequently lost for each. We then add this with the other 51% of free throws which do have a chance to be rebounded at a 14% clip. Each one of these misses that is not rebounded also comes with a half of possession lost.

So for example, if you missed 2 free throws, which many coaches feel like ends up being a loss of a full possession, we see that the formula calculates

= [2 * 0.5 * 0.49] + [2 * 0.5* (1-0.14) * 0.51]

= 0.9286

a loss of 0.9286 of a possession for every missed pair of free throws.

The Formula

When we put it all together, we can finally determine a formula to calculate Points Per Lost Possession (PPLP):

$$PPLP - Equation\ (1)$$

$$= \frac{\mathrm{PTS}}{[\mathrm{TO's}] + [\mathrm{FGM} * (1 - \mathrm{ORB\%})] + [(\mathrm{FTM} * 0.5 * 0.49) + (\mathrm{FTM} * 0.5 * 0.51 * (1 - 0.14))]}$$

With this proposed new metric, how can we show that acquiring NBA players using PPLP as one measure is more effective than other proven metrics, most notably a metric that has been used for more than a century?

Chapter 3: Data Analysis and Results

The Data

The NBA team data used in the analysis was obtained from Basketball Reference.com. The data consisted of the following variables for the seasons 2004 through 2012: field goals made, field goals attempted, three-pointers made, three-pointers attempted, free throws made, free throws attempted, turnovers and points. Using this dataset, we applied Equation (1) to calculate each team's season PPLP.

Defining Success

In order to compare PPLP against any another metric, most notably points scored, we needed to define team success in the NBA. For the purpose of this paper, we defined success as a team qualifying for the NBA playoffs. With NBA team success now defined, we can compare the new PPLP metric previous standard of points scored. We will also include a comparison of PPLP to a widely used forecasting system called Accuscore.

Logistic Regression

Logistic regression analysis is a useful tool for analyzing predictive models that contain a binary response variable. Our response variable is playoff qualification, which is binary, and is coded (1) for making the playoffs and (0) for not making the playoffs. This will be predicted by two separate explanatory variables, points scored and PPLP. We are intending to model the probability of making the playoffs as explained by points and PPLP. Because we are interested in determining how well each of the unique predictors would be tied to playoff qualification we will use a logit link model for our analysis.

In a logistic regression analysis, our null hypothesis is always that the coefficient for the predictor or explanatory variable equals zero and does not have a significant effect on predicting the response variable. The alternative is that the explanatory variable's coefficient is not zero, and does in fact significantly predict the response variable. Of course, for this analysis we hope to see that PPLP is indeed a significant predictor of playoff qualification, and if it is, that it is also a stronger predictor than points scored.

Using Minitab (version 16), we found PPLP to be a significant predictor of playoff qualification (p-value less than 0.0001). PPLP was also a stronger predictor of playoff qualification than points scored (Z-statistic of 7.15 and 78.3% concordant) compared to points scored (Z-statistic of 2.35 and 62.4% concordant). Percent concordant measures all possible 0,1 pairs of data to determine which percentage of the paired observations would be classified correctly based on the observed PPLP values. Table 4-1 in the appendix provides a summary of the results.

Therefore, the evidence suggests that PPLP outperforms points scored when used as the only predictor in a model predicting whether or not a team qualifies for the NBA playoffs. These initial results are promising, and help to validate using PPLP as an evaluation statistic instead of simply points scored. But one final question remains. Does PPLP work when applied to the NBA as it currently exists?

Chapter 4: Application to Previous NBA Seasons

The Model

In order to apply PPLP as a predictor of playoff qualification, we begin with the logistic regression model our data produced:

$$Probability\ of\ Making\ the\ Playoffs - Equation\ (2)$$

$$= \frac{1}{1 + \mathrm{e}^\wedge - (-21.7655\ +\ 10.991 * \mathrm{PPLP})}$$

From this model, we can forecast the probability of any one NBA team making the playoffs by plugging in their projected PPLP. Table 4-2 summarizes a few benchmarks across the domain of PPLP scores.

To examine the relevancy of our model, we will look at a series of case studies from the current 2013-14 NBA season, and compare the PPLP forecast of a team's playoff chances against a respected and frequently used forecasting resource. The final results, how the teams have turned out present day, could potentially provide better insights on the practicality and viability of the PPLP application in the NBA as a whole.

Accuscore

Accuscore, is a professional forecasting system which is used to predict outcomes of nearly every professional sporting event. In the world of sport's statistics, Accuscore forecasts are used widely. Some of their most notable clients include ESPN, CBS, and Yahoo! Sports. For a standard user to gain access to the full breadth of Accuscore predictions and forecasts, the cost is $300 per person per year. Because of Accuscore's far-reaching and proven acceptance in the world of sports, we will use their projections as a benchmark in our case studies to compare against similar forecasts from our PPLP model (AccuScore Advantage).

Case Study #1- New York Knicks

The New York Knicks, led by Carmelo Anthony were coming off of a strong 54-win season in 2012-13, and nearly every NBA expert felt that they were a sure thing to make the playoffs in 2013-14. Accuscore's forecast matched the experts, and on October 28th, 2013 Accuscore released a prediction in which their calculations forecasted the New York Knicks making the postseason with a 90.9% probability—a near certainty (NBA Season Predictions). However, when examining the team's projected PPLP, our model would arrive at a less certain forecast.

Using the New York Knicks' players who stood to play the most minutes at the start of the 2013-14 we took a weighted average of each individual player's historical PPLP (over the past 3 seasons) to determine a team projected PPLP for the 2013-14 Knicks. With that team's projected PPLP, we gave our own forecast of the Knicks' playoff probability. A breakdown of each individual Knick player's PPLP, and a projected team PPLP for the 2013-14 season[2] can be found in Table 4-3.

[2]Note: we did not forecast a large quantity of playing time for rookie shooting guard, Tim Hardaway Jr., and thusly, he was left out of these calculations.

From the table we see that our forecasted PPLP for the Knicks current 2013-14 season was 2.08. Applying Equation (2), an NBA team with a 2.08 PPLP has a 0.749 probability of making the playoffs. The PPLP model predicts the Knicks to have a reasonable chance of qualifying for the playoffs—roughly a 7 out of 10 chance, which is significantly less than the 91% offered by Accuscore.

How did this current season turn out? At present, the Knicks stand at 29-41 on the year. They are three games out of the playoffs with just two weeks remaining in the regular season. To the surprise of most NBA experts, it seems as if the Knicks will indeed miss out on qualifying for the playoffs. However, our PPLP model forecasted a much less optimistic outlook for the Knicks than what was the general consensus, and we are not nearly as surprised by their disappointing season.

Case Study #2- Detroit Pistons

The Detroit Pistons are a basketball statistician's evil nemesis. Their public bashing of analytics and cold insistence that they will only evaluate basketball talent based on what meets the eye, leaves basketball data analysts exasperated. We don't think that our numbers provide all the answers. But neither do your eyes. Piece them together, and you can end up with something special. For the Pistons' front office to completely ignore data that could provide valuable insights is baffling. Yet, this is the Pistons model, and this is how they continue to operate (Mayo). It has fooled their fans, and even some experts who follow basketball.

We take this current NBA season for example. The Pistons had an active offseason, acquiring flashy and high priced players like Josh Smith and Brandon

Jennings. Pairing those two players with big men Greg Monroe and Andre Drummond was supposed to make the Pistons legitimate playoff contenders in the NBA Eastern Conference. Heralded sports analysts, Jalen Rose and Bill Simmons ranked them as the 12^{th} best team in the NBA this preseason. Early in the season, ESPN ranked them 14^{th} overall in the NBA, while Yahoo! Sports ranked them at 10^{th} and SB Nation ranked them 13^{th}. All of the expert analysts predicted them to make the playoffs. Certainly, the Pistons front office figured that their no-data management style was set up to work beautifully. Running the Pistons projected lineup through our PPLP model[3] Equation (2) yielded a strikingly different forecast. The detailed player-by-player Piston's PPLP summary can be found in Table 4-4.

What our results found was a projected 2013-14 PPLP for the Detriot Pistons of 1.97. Plugging that number into Equation (2), we forecasted the Detroit Pistons to have a 47.1% chance of making the playoffs at the start of the season. This means they were more likely to miss the playoffs than make them. Their two biggest and most expensive offseason signings, Josh Smith and Brandon Jennings, carry with them poor PPLP scores that hinder Detroit's chances of reaching the playoffs.

As it turned out, the Detroit Pistons currently stand at 25-44 on the season, and are all but mathematically eliminated from qualifying for the post-season. While this is shocking news to Detroit Pistons fans, and a high proportion of the NBA expert community, the PPLP model did not forecast a promising season for the Detroit Pistons.

[3]Note: since Kentavious Caldwell-Pope was a rookie, we projected his PPLP using CBS.com preseason projections.

As the two case studies indicated, the PPLP model can be a useful forecasting tool when applied to the NBA today.

Chapter 5: Concluding Remarks

A Vision for Future Work

What have we accomplished with PPLP? There is evidence supporting that PPLP could be a better predictor of success in the NBA as measured by playoff qualification than points scored. On top of that, the PPLP model can work when applied to the current structure of the NBA.

The long-term vision for a PPLP model would be to construct a complete NBA player database, in which their PPLP would be calculated based on their most recent season(s) performance. With PPLP data on every player in the NBA market, one could expose inefficiencies, and sign players that the PPLP model values more than the rest of the league. These would be the building blocks for developing league-leading analytics in the NBA, analogous to when the Oakland Athletics evaluated baseball players based on OBP instead of batting average. Perhaps one day, PPLP will be regularly reported in box scores, and player's season stat lines in place of, or alongside, points scored.

Constructing the Dream Team

As a further exercise, we used PPLP to construct a "Dream Team" that could potentially be acquired by the 2015-16 NBA Season. This team would fall under the NBA Salary cap, and could lead to grand success[4]. A breakdown of the team construction can be found in Table 4-5.

[4] A caveat: my dream team "drafted" Kenneth Faried.

With a projected PPLP of 2.27, the model estimates that this team would have a 96% chance of making the playoffs. Making the playoffs 96 years out of 100 would be great success, and potentially provide for an avid fan-base. Success at the highest level of basketball could yield far-reaching financial returns, and provide monetary validation for the PPLP model.

A Final Caveat

It is important to realize that the PPLP model does not account for how the players selected would interact together once placed together. Parts of player interaction can be vital to overall team success such as: team chemistry, locker room personalities, ego management, etc. PPLP does not forecast the outcome of these intangibles. In order to achieve highest success, a better suited method could be to combine both the observations of player interaction from a human perspective and the PPLP model.

APPENDIX

Table 4-1- **Comparing PPLP to Points Scored**

1. Binary Logistic Regression: Playoffs? versus PTS

[Current Metric used in NBA]	
Z-Value:	2.35
P-Value:	0.016
Percent Concordant:	62.4%

2. Binary Logistic Regression: Playoffs? versus PPLP

[Invented Metric]	
Z-Value:	7.15
P-Value:	0.000
Percent Concordant:	78.3%

Table 4-2- **Playoff Probabilities**

PPLP	Probability of Making Playofffs
2.50	99.67%
2.40	99.02%
2.30	97.11%
2.20	91.79%
2.10	78.84%
2.00	55.39%
1.90	29.26%
1.80	12.11%
1.70	4.39%
1.60	1.50%
1.50	0.05%

Table 4-3- **New York Knicks PPLP**

	Projected MPG	PPLP
Anthony, Carmelo	36	2.26
Bargnani, Andrea	30	2.08
Smith, JR	30	2.07
Felton, Raymond	30	1.68
Chandler, Tyson	25	3.10
Shumpert, Iman	25	1.67
Stoudemire, Amar'e	20	2.33
Martin, Kenyon	20	1.93
Prigioni, Pablo	20	1.57
Team TOTAL		2.085212

Table 4-4- **Detroit Pistons PPLP**

	Projected MPG	PPLP
Smith, Josh	35	1.87
Jennings, Brandon	35	1.84
Monroe, Greg	30	2.10
Drummond, Andre	30	2.36
Stuckey, Rodney	25	2.00
Singler, Kyle	25	1.90
Bynum, Will	20	1.76
Caldwell-Pope, Kentavious	20	1.82
Jerebko, Jonas	15	2.16
Team TOTAL		1.979362

Table 4-5- **The Dream Team**

	Projected MPG	PPLP	Salary
Young, Thaddeus	35	2.31	$10 million
Jefferson, Al	32	2.42	$13 million
Millsap, Paul	30	2.20	$9 million
Afflalo, Arron	30	2.28	$7 million
Dragic, Goran	28	2.10	$7 million
Redick, JJ	25	2.28	$6 million
Faried, Kenneth	25	2.50	$2 million (rookie contract)
Collison, Darren	20	2.04	$2 million
Patterson, Patrick	15	2.30	$3 million
Team TOTAL		2.276208	

Figure 1-1

Binary Logistic Regression: Playoffs? versus PTS

```
Link Function: Logit


Response Information

Variable   Value  Count
Playoffs?  1        144  (Event)
           0        126
           Total    270


Logistic Regression Table

                                                 Odds     95% CI
Predictor       Coef    SE Coef      Z       P  Ratio  Lower  Upper
Constant    -3.48592    1.54353  -2.26  0.024
PTS        0.0004575  0.0001943   2.35  0.019   1.00   1.00   1.00


Log-Likelihood = -183.657
Test that all slopes are zero: G = 5.786, DF = 1, P-Value = 0.016


Goodness-of-Fit Tests

Method           Chi-Square   DF      P
Pearson             244.895  250  0.579
Deviance            332.995  250  0.000
Hosmer-Lemeshow      10.996    8  0.202


Table of Observed and Expected Frequencies:
(See Hosmer-Lemeshow Test for the Pearson Chi-Square Statistic)

                                  Group
Value     1     2     3     4     5     6     7     8     9    10  Total
1
  Obs    13    11    12    11    11    16    15    19    19    17    144
  Exp   9.6  13.1  14.4  14.8  14.6  14.9  15.2  15.5  16.0  15.9
0
  Obs    14    16    16    17    16    11    12     8     8     8    126
  Exp  17.4  13.9  13.6  13.2  12.4  12.1  11.8  11.5  11.0   9.1
Total    27    27    28    28    27    27    27    27    27    25    270


Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant   11324     62.4  Somers' D              0.26
Discordant    6665     36.7  Goodman-Kruskal Gamma  0.26
Ties           155      0.9  Kendall's Tau-a        0.13
Total        18144    100.0
```

Figure 1-2

Binary Logistic Regression: Playoffs? versus PPLP

Link Function: Logit

Response Information

Variable	Value	Count	
Playoffs?	1	144	(Event)
	0	126	
	Total	270	

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Constant	-21.7655	3.05507	-7.12	0.000			
PPLP	10.9910	1.53639	7.15	0.000	59336.87	2921.08	1205328.80

Log-Likelihood = -150.046
Test that all slopes are zero: G = 73.008, DF = 1, P-Value = 0.000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	269.711	268	0.459
Deviance	300.091	268	0.086
Hosmer-Lemeshow	4.619	8	0.797

Table of Observed and Expected Frequencies:
(See Hosmer-Lemeshow Test for the Pearson Chi-Square Statistic)

	Group										
Value	1	2	3	4	5	6	7	8	9	10	Total
1											
Obs	3	8	10	12	12	12	20	19	23	25	144
Exp	3.8	6.8	8.7	10.8	13.2	15.5	17.3	20.3	22.8	24.7	
0											
Obs	24	19	17	15	15	15	7	8	4	2	126
Exp	23.2	20.2	18.3	16.2	13.8	11.5	9.7	6.7	4.2	2.3	
Total	27	27	27	27	27	27	27	27	27	27	270

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	14211	78.3	Somers' D	0.57
Discordant	3882	21.4	Goodman-Kruskal Gamma	0.57
Ties	51	0.3	Kendall's Tau-a	0.28
Total	18144	100.0		

BIBLIOGRAPHY

"AccuScore Advantage." *AccuSore*. N.p., n.d. Web. 22 Mar. 2014. <https://accuscore.com/join-now>.

Berri, Brook and Schmidt. "Does one need to Score to Score?" *International Journal of Sports Finance.* (2007), pp. 142-148. <http://www.suu.edu/faculty/berri/ScoretoScoreI JSF.pdf>

"Bill James." *BR Bullpen*. N.p., n.d. Web. 22 Dec. 2013. <http://www.baseball-reference.com/bullpen/Bill_James>.

"Dean Oliver's Four Factors." *NBA.com*. N.p., n.d. Web. 23 Mar. 2014. <http://www.nba.com/thunder/news/factors050127.html>.

"Four Factors - Team Evaluation Metrics." *NBAstuffer.com*. N.p., n.d. Web. 13 Dec. 2013. <http://www.nbastuffer.com/component/option%2Ccom_glossary/Itemid%2C90/catid%2C41/func%2Cview/term%2CFour%2520Factors/>.

George, Rodman. "Interview of George Rodman." Telephone interview. Dec. 2012.

Jensen, Shane. "Measuring Player Contributions in Hockey." *Chance*. N.p., 2013. Web. 24 Nov. 2013.

Lewis, Michael. "The No Stats All Star." The New York Times. The New York Times Company, 15 Feb. 2009. Web. 18 Dec. 2012.

Mason, Jeffrey. "Salary of NBA Players: A Regression Analysis." *Geocites.* < http://www.geocities.ws/jeffreylylemason/Economics/NBA.pdf>.

Maymin, Phillip. "Acceleration in the NBA: Towards an Algorithmic Taxonomy of Basketball Plays." Sloan Sports Conference. ESPN, 1 Mar. 2013. Web. 14 Mar.

2013.<http://www.sloansportsconference.com/wpcontent/uploads/2013/Acceleration%20in%20the%20NBA%20Towards%20an%20Algorithmic%20Taxonomy%20of%20Basketball%20Plays.pdf>.

Mayo, David. "How Deeply Have the Detroit Pistons Delved into Advanced Stats?" *Michigan Live*. N.p., n.d. Web. 24 Mar. 2014. <http://www.mlive.com/pistons/index.ssf/2014/01/ask_david_how_deeply_have_detr.html>.

Morris, Jerod. "In Honor of Dirk: 'Points Per Miss' and His Spot among the NBA's All-time Greats." Midwest Sports Fans. N.p., 11 May 2011. Web. 08 May 2013. <http://www.midwestsportsfans.com/2011/05/in-honor-of-dirk-points-per-miss/>.

"NBA Basketball." *TeamRankings.com*. N.p., n.d. Web. 3 Jan. 2014. <http://www.teamrankings.com/nba/>.

"NBA Game Charting: Floor Locations from 82games.com." *NBA Game Charting: Floor Locations from 82games.com*. N.p., n.d. Web. 23 Dec. 2013. <http://www.82games.com/locations.htm>.

"NBA Offensive Rebuilding Studies from 82games.com." *NBA Offensive Rebuilding Studies from 82games.com*. N.p., n.d. Web. 12 Nov. 2013. <http://www.82games.com/comm13.htm>.

"NBA Season Predictions: AccuScore Data." *Metta Chronicles*. N.p., n.d. Web. 1 Mar. 2014. <http://mettachronicles.com/nba-season-predictions-accuscore-data/>.

Packey, Brian. "2013 NBA Power Rankings: Detroit Pistons Moving up." *SB Nation*. N.p., n.d. Web. 24 Mar. 2014. <http://www.detroitbadboys.com/2013/11/4/5065674/2013-nba-power-rankings-detroit-pistons>.

Sathe, Mihir. “Basketball: Regression Model to Correlate Performance with Salary.” *University of Southern California.* < http://www.isi.edu/~szekely/contents/cs548/2013-spring/game7.pdf>

Smeltz, Nate. "How Hollinger Crunches the NBA." *ESPN Front Row*. N.p., 23 Mar. 2011. Web. 2 Mar. 2014. <http://frontrow.espn.go.com/2011/05/how-hollinger-crunches-the-nba/#more-5572>.

Smith, Sekou. "2013-14 NBA Salary Cap Figure." *Hang Time Blog with Sekou Smith at NBAcom*. N.p., 9 July 2013. Web. 12 Nov. 2013. <http://hangtime.blogs.nba.com/2013/07/09/2013-14-nba-salary-cap-figure-set-at-58-679-million/>.

Stein, Marc. "Grizzlies Hire John Hollinger." *ESPN*. ESPN Internet Ventures, 14 Dec. 2012. Web. 18 Dec. 2013. <http://espn.go.com/nba/story/_/id/8745158/espn-columnist-john-hollinger-join-memphis-grizzlies-front-office>.

Surowiecki, James. "A Conversation with Bill James." *Slate Magazine*. N.p., n.d. Web. 21 Feb. 2014. <http://www.slate.com/articles/arts/culturebox/2003/06/moneyball_redux.html>.

"The Business of Sports." *Top Business Degrees*. N.p., n.d. Web. 23 Jan. 2014. <http://www.top-business-degrees.net/sports/>.

"The Man Behind The 'Moneyball' Sabermetrics." *NPR*. NPR, n.d. Web. 14 Jan. 2014. <http://www.npr.org/2011/09/26/140813409/the-man-behind-the-moneyball-sabermetrics>.

Van Gundy, Stan, Kevin Pritchard, Mike Zarren, and RC Buford. "Basketball Analytics 2013." 2013 MIT Sloan Analytics Conference. Boston Convention Center, Boston, MA. 3 Mar. 2013. Speech.

Weiss, Eric, and Kirk Goldsberry. "The Dwight Effect: A New Ensemble of Interior Defense Analytics for the NBA." Sloan Sports Conference. ESPN, 1 Mar. 2013. Web. 9 Mar. 2013. <http://www.sloansportsconference.com/wp-content/uploads/2013/The%20Dwight%20Effect%20A%20New%20Ensemble%20of%20Interior%20Defense%20Analytics%20for%20the%20NBA.pdf>.

Wagner, Ellen. "Data Changes Everything: Delivering on the Promise of Learning Analytics in Higher Education." *Educase*. N.p., n.d. Web. 19 Dec. 2013. <http://www.educause.edu/ero/article/data-changes-everything-delivering-promise-learning-analytics-higher-education>.

"2013-14 NBA Player Contracts." *Basketball-Reference.com*. N.p., n.d. Web. 23 Feb. 2014. <http://www.basketball-reference.com/contracts/players.html>.

ACADEMIC VITA

Derek Gerberich
derek.gerberich@gmail.com

EDUCATION	**Pennsylvania State University- The Graduate School** -Master's Degree in Applied Statistics	May 2014
	Pennsylvania State University- Schreyer Honors College -Bachelors of Science in Applied Statistics -Minor in Communications Arts and Sciences	May 2014
EXPERIENCE	**National Center for Education Statistics (NCES)- Research Analyst** -Analyzed student response-patterns to determine reasons for incomplete assessments -Created Access database to track U.S. product publication for international assessments	2013
	American Institutes for Research- Research Associate -Worked as a consultant while jointly employed at NCES -Designed Excel Macros to expedite analysis of international country profiles	2013
	Penn State Men's Basketball- Director of Basketball Analytics -Innovate and develop statistical metrics to improve team -Responsible for working up to 20-25 hours per week	2010-present
	Lancaster Barnstormers- Baseball Information Analyst -Produced a statistical analysis addressing issues slowing down the average game time	2011
	BleacherReport.com- Featured Columnist for the Dallas Cowboys -Published over 70 articles totaling over 300,000 pageviews -Promoted to Featured position after beginning as an intern Summer 2012	2012
	Midwestsportsfans.com- Head Fantasy Baseball Writer -Hired and directed staff of 6 staff writers responsible for producing season-long content	2010-2012
	Lebanon Valley Midget Baseball League -Umpire of nine years	2005-2013
ACTIVITIES	**Penn State Cross Country Club** -Four-Time NIRCA National Championship Qualifier	2010-present
	Alliance Christian Fellowship (ACF) -Named ACF Outreach Elder for 2012-2014 -Lead executive board, and direct the trajectory of campus-wide ministry	2010-present
PUBLICATIONS	**(Links to feature articles from Bleacher Report and Midwest Sports Fans)**	
	10 Reasons Dallas Cowboys Fans Have to Be Excited for 2012 Season- **September 3rd, 2012** Penn State Football Scandal: Why NCAA Punishments are Unreasonable- **July 26th, 2012** Connecting all 344 NCAA Basketball Teams by the Transitive Property- **March 5th, 2012**	
AWARDS	JPSM Junior Fellow (2013)	
SKILLS	**SAS, Pivot Tables, Simulations, Excel, Access, SQL, R, Minitab, PowerPoint, C++, SharePoint**	