New Approaches to Player Valuation: Analyzing How Wins Generate Revenue for Major League Baseball Teams

W. Graham Tyler
Brown University 2012

Economic theory suggests that correct valuation of free agent players in Major League Baseball requires estimation of the revenue generated by the added wins the player contributes through his on field performance. Close analysis of the research to date in this area, however, suggests flaws in previous estimation techniques analyzing the revenue-win relationship, which lead to biased estimates with significant implications for estimating a player's worth to a team. This paper departs from past work by estimating team specific, nonlinear effects of winning on revenue and controlling for the effect price has on this relationship. This analysis finds significant variation across teams in the revenue-win relationship, much of which is not explained by market demographic factors and almost none of which is explained by overall market size. Additional analysis suggests that it is difficult to even group seemingly similar teams based on their revenue-win relationship, implying that the returns to winning are determined by the unique, dynamic relationship each team has with its fans and other factors that cannot be quantified in a model that pools all teams.

The author wishes to especially thank Professors Ken Chay and John Tyler for advising the research process and Liz Tyler and Maurine Tobin for comments on earlier drafts of this paper. Additionally, he would like to thank Steve August, Jon Budish, John Burger, Dan Barbarisi, Brian Cashman, Dan Duquette, Michael Fishman, Vince Gennaro, Stephen Walters, and various sources with experience in Major League Baseball for insight that helped to frame the research and analysis contained in this paper.

The advent of free agency in 1975 fundamentally changed the economics of Major League Baseball (MLB). Previously indentured to a team for their entire career and forced to accept whatever salary owners felt fair, players now had the ability to move from team to team, creating a competitive market for their services. The result was an exponential increase in salaries, culminating with Alex Rodriguez's famous ten year $252 million contract with the Texas Rangers in 2000 (Verducci, 2000). The increased payroll spending did not occur uniformly across teams, however, and in 2000 the disparity in the payrolls and success between the top teams and the bottom teams led to the creation of the Blue Ribbon Panel hired by MLB to investigate the financial sustainability of the league. The panel's findings—that differences between teams in local revenue severely hindered competitive balance—implied what fans already knew: that in the free agent market some teams had the ability to vastly outspend their competition.

The creation of free agency and its subsequent effect on salaries and competitive balance led naturally to two fundamental questions with implications for baseball teams, fans, and economists that are implicitly discussed in the media, are explicitly discussed every day by baseball executives, and that have been empirically analyzed by economists. The critical questions to ask about the free agent market are how much a player is truly worth and what determines this value. Economic theory suggests that teams should not pay a player more than the added revenue generated by signing the player to the team; the implication of this theory is that the disparities between teams in free agent spending are not due to differences in baseline revenue, but rather to differences between teams in the marginal revenue from winning. Winning, therefore, must be more valuable to the Yankees than to anyone else, given that they consistently outspend their competition in the free agent market.

The uniquely statistical nature of baseball makes player valuation a popular topic among many economists and fans. Almost every relevant action in the game is recorded as an individual feat, and has been for most of MLB's history. The result is a mass of data on player performance that statisticians have been analyzing since the 1970's in order to estimate a player's contribution to the team's success, independent of the performance of those around him. Unlike a teacher or a doctor, for whom the marginal product of labor is difficult to quantify, if we consider a baseball player's product to be added wins, we can theoretically quantify the marginal impact of his performance on the team's output by analyzing this performance data. This realization led to Gerald Scully's seminal work in 1974 that attempted to estimate players' marginal revenue product (MRP), and has spawned ongoing academic and commercial research into how to apply such estimations. Close reading of this literature, however, reveals both inconsistent findings between estimation techniques and potential confounding effects that could bias the estimates.

The goal of this paper is to explore these issues, analyze how they may have biased previous work, and incorporate some previously unconsidered factors that may determine the added revenue from winning, in order to estimate a player's value. First, I will analyze the extent to which the models to date are potentially misspecified. Second, I will address factors not previously considered that may have profound effects on our understanding of a team's revenue return to winning. Finally, I will suggest additional considerations teams should make when estimating player value that are difficult to capture in a statistical model for revenue.

As Berri and Schmidt (2006) find, over the last few decades fans have grown increasingly responsive to the competitiveness of their team, increasing the returns to winning and the money that teams stand to make from efficient decision making. Additionally, the rise in salaries from the advent of free agency has also made effective player valuation significantly more important to teams' financial and competitive success. This suggests that a better understanding of how teams should spend the millions of dollars they devote each year to player salaries within the context of the free agent market is ever more relevant, but the implications of the research also stretch beyond baseball.

To the extent that we can reliably estimate a worker's MRP, this estimation can allow us to analyze various economic theories and topics. Estimates of a baseball player's MRP, for example, have been used to study monopsonistic exploitation (Raimondo, 1983), owner collusion in driving down salaries (Bruggink and Rose, 1990), the presence of a winner's curse in auctions (Cassing and Douglas, 1980), and human capital investment theory (Blass, 1992). Additionally, MRP estimates have been used to look at methods within baseball to increase competitive balance (Zimbalist, 1992 and 2001; Burger and

Walters, 2003) and to develop strategies for efficiently building a winning and financially successful franchise (Gennaro, 2007; Silver, 2006). Anecdotal evidence from talking to current and former figures working in or around major league franchises suggests, however, that few teams utilize estimates of MRP when making player personnel decisions (August, Interview; Barbarisi, Interview; Cashman, Interview; Duquette, Interview). This indicates either that owners are not the profit maximizing agents economists would assume them to be, or that the decision making process within the industry is inherently flawed and without perfect information.

The rest of this paper will proceed as follows: Section 1 will detail the history of MRP estimation for Major League Baseball players both by academics and commercial authors. Sections 2 and 3 will outline my hypotheses about the nature of the revenue-win relationship for MLB teams, the subsequent implications for MRP estimation, and my strategy for testing these hypotheses. Section 4 details my data and the results of my analysis. Section 5 discusses the implications of my findings and possible theoretical explanations for my results. Section 6 provides the potential caveats to my analysis and explains why they may or may not be significant. Section 7 concludes and details potential areas of further research on the topic. At the end of the paper there is an appendix containing significant tables and figures that will be discussed in Sections 4, 5, and 6.

I. History of MRP Estimation Strategy

MRP estimation strategy for MLB players falls into three categories that I will call the Krautmann model, the commercial model, and the Scully model. The Krautmann model, developed by Anthony Krautmann (Krautmann, 1999), uses the classic economic theory of markets to derive a player's MRP. Krautmann assumes that free agency creates a free market in which teams reveal the marginal revenue generated by a player at a certain position through the salary they pay that player. This salary then represents the marginal revenue product of all similarly performing players at that position. The drawback to this model is that it makes very strong assumptions about players' value to teams. First, it assumes that each team values players the same, a phenomenon that is refuted by the differences across baseball in team payrolls. Second, it assumes that the return to a win is linear, so that a team that will win

95 games by signing a given free agent will derive the same value as a team that will win 75 games if they add the free agent. Third, it assumes that teams are privy to perfect information about how a player will perform and about how this performance will impact their bottom line. Finally, it assumes that the free agent market is perfectly competitive so that any given team must pay exactly a player's MRP, or the player will sign with another team.

The commercial model utilized by analysts Vince Gennaro in *Diamond Dollars* and Nate Silver in *Baseball Between the Numbers* seeks to construct the revenue-win relationship through the authors' knowledge of how baseball teams generate revenue. Utilizing various estimation methods, these authors fit models relating ballpark attendance to wins to estimate general fan responsiveness to winning, either uniquely for teams in the case of Gennaro or generally across teams in the case of Silver. They then multiply these estimates by average ticket price to get attendance revenue at each win total. Next, they assume the revenue generated from concessions, merchandise, parking, and box seats is a fixed proportion of the revenue generated by attendance. Multiplying the attendance revenue at each win total by the set proportion then yields the added revenue from these factors. Finally, they allow for a nonlinear relationship between revenue and wins by estimating the expected bump in revenue from making the playoffs and multiplying this by the probability of making the playoffs at each win total. While these models assume a nonlinear relationship between revenue and wins—and, in the case of Gennaro, assume team specificity—they employ assumptions about how attendance revenue is extrapolated to total revenue. Additionally, construction of the revenue-win relationship in this manner makes it hard to quantify and analyze the variation between teams in the returns to winning as I intend to.

The Scully model is built upon the two step process conceived by Gerald Scully in 1974 and is the most compelling for empirically analyzing the revenue-win relationship for teams within the framework of economic theory. Scully posits that a player's MRP is best modeled as the additional wins he is responsible for, multiplied by the marginal revenue associated with these wins. Scully, and others utilizing his approach, fit models with revenue as the dependent variable and wins as the explanatory variable of interest, controlling for other factors affecting revenue. Then they estimate a player's added wins based on performance data and multiply these estimated wins by the estimated coefficient on wins from the revenue-wins model. In this approach, a player's MRP is a linear function of his performance and is not specific to each team, implying that every team should be willing to spend the same amount for a given player.

The prominent academic hypothesis refuting players' equal value across teams is that a team's market size impacts their marginal return to wins. Models taking this into account interact a measure of market size with team wins in the revenue regression to allow for a differential effect of wins on revenue depending on the size of the market within which a team operates. This strategy has found conflicting results, with Sommers and Quinton (1982) finding that the returns to wins are greater in larger markets and Zimbalist (1992) finding no significant effect of market size.

More recently, Burger and Walters (2003) added complexity to the market size model with their theory that each market is made up of either purist fans or bandwagon fans. Their model assumes the proportion of each type of fan is the same across markets, but larger markets will have more of each and smaller markets will have less of each. Bandwagon fans only become interested in a team and create revenue when a team hits a threshold number of wins. Purists generate revenue at any number of wins, but will become more intense and therefore create more revenue once the team hits the threshold. Burger and Walters estimate a nonlinear revenue-wins relationship using a spline function at the wins threshold, and find that "teams in the largest markets derive up to six times more marginal revenue from each additional victory than teams in the smallest markets" (Burger and Walters, 2003, p.121).

The work of Burger and Walters is the most compelling of the Scully-type models, as it allows for both team specificity and nonlinearity in the revenue-wins relationship. There are, however, questions that arise within their work based on the assumptions that they make. The first caveat is the assumption that market size is the only team specific factor that influences the marginal revenue of a win. It is plausible that there are large market teams for whom wins may not have a significant impact on revenue, even in the current era. For example, the Cubs, who reside in Chicago—one of baseball's largest markets—have a long history of losing that they have successfully marketed to their fans. The marketing has worked so well that Cubs fans seem to enjoy flocking to Wrigley Field to watch the “lovable losers” fail to win baseball games. In Miami, the combination of a rainy climate and an outdoor stadium seem to have the opposite effect on Marlins fans—they rarely attend games, even during winning seasons (although it remains to be seen whether a brand new stadium with a retractable roof will change this). This suggests that the hypothesis that every market is made up of the same proportion of purist and bandwagon fans is not one to be accepted without deep analysis. Instead, this anecdotal evidence would suggest that there may be hard to quantify, very team specific factors that determine the relationship between teams and their fans.

The second caveat is that, as Vince Gennaro and countless others find, the probability of making the playoffs at a given win total is nonlinear. Instead it resembles a standard probit curve, with the largest marginal gain in probability at those win totals right on the cusp of making the postseason. This suggests that, though the spline function utilized by Burger and Walters addresses the nonlinearity of the relationship and correctly attributes it to the possibility of making the playoffs, the revenue-win relationship above the threshold should be nonlinear. Additionally, one issue not addressed in the literature is the fact that the probability of making the playoffs at a given win total each year is historically different depending on the division a team plays in. Since the three divisions per league format was instituted in 1995, the American League East with the Yankees and Red Sox has been notoriously difficult to win, meaning teams need more wins to reach the postseason. On the other hand, the American League Central is historically mediocre, requiring fewer wins to reach the postseason. If all that fans care about is making the playoffs, then fans of teams in the AL Central should require fewer wins to “jump on the bandwagon”.

Finally, few papers attempt to account for the effect that ticket price has on demand, a counterintuitive concept given that we are essentially estimating a demand curve, with demand being a function of the quality of the product. One needs only look at the vast differences in ticket prices between teams to understand that, given the distance between teams and the unique relationship each team has with fans, franchises are price setting monopolies. If owners understand that demand is related to the success of the team, they should take this into account when setting prices. Teams have the privilege of observing their players on a daily basis before the season and should have the best idea of how they will perform, meaning their predictions for performance should be closely correlated with actual performance. If this is the case, then ticket price should be positively correlated with actual wins and with revenue, meaning that price would be an omitted variable.

II. Hypotheses and Assumptions

In order to more fully understand how wins generate revenue for MLB teams and subsequently how much a player is worth I will estimate two separate types of models to test three different hypotheses. The first model will estimate how much revenue a win generates for every team and will be based on two assumptions about the revenue-win relationship. This model will also allow me to test whether or not price is an omitted variable. The second model will test how much variation in the revenue-win relationship can be explained by observable market demographic factors that are relatively stable over time.

The first assumption is that the marginal revenue of a win varies significantly across teams. This means that two teams both going from 85 wins to 86 wins in a season, for example, will generate very different amounts of revenue from this extra win. While this seems like a straightforward assumption, it is nonetheless the case that the bulk of the research estimating players' MRP fits a model that pools teams to create an across teams estimate for the value of a marginal win. Given that a player's value is derived from the added revenue a team generates from the player's performance, the assumption implies that the same player will have very different values to different teams. Therefore, to the extent that teams act on the anticipated value of the player and players respond rationally, we should see players systematically signing with the team that values them the most. If players have different true values to different teams, then it is not necessarily the case that the winning bidder for a given free agent overbid, as the theory of the winner's curse in auctions would predict. In fact, as Burger and Walters (2003) point out, a team can bid well under a player's MRP for their franchise and outbid their competition if they are bidding against a team that values the player much less than they do.

The second assumption is that the revenue-win relationship is nonlinear. More recent work generally accepts this hypothesis to be true, but much of the older research on player MRP estimates a linear model relating revenue and wins, despite the intuitive argument that not all wins are created equal. It is unlikely that going from 65 wins to 66 generates the same fan response and subsequent revenue as going from 91 to 92 wins. In the latter case, the additional win raises the probability of making the playoffs, generating more fan excitement and increasing the chances of additional gate revenue during the current season from playoff games. The differential effects on revenue may be even greater the following season, as making the playoffs would likely increase demand in the form of more ticket sales or consumer absorption of increases in prices; between 1993 and 2006 teams that reached the postseason raised ticket prices the following year by about 10%, while teams that failed to make the playoffs raised them by less than 6% (Gennaro, 2007).

The first hypothesis (H1) I will test is that price is an omitted variable in specifications of the revenue-win model when not included as an independent variable. As explained above, if price is determined endogenously based on predictions of the team's wins and these predictions are closely correlated with the team's actual wins, then price is an omitted variable. To understand the implications of the potential bias we must understand the extent to which price is correlated with predicted wins. A poorly performing team should face reduced demand and thus diminished revenue. If this poor performance is unexpected then this effect will be coupled with high prices set in anticipation of a good team, which will further decrease attendance and revenue. Similarly, setting prices low and then unexpectedly having a good team would deflate some of the added revenue from the team's success. If there is a lot of uncertainty even by owners as to how a team will do, then they should set price somewhere in the middle, in which case there is little correlation between price and predicted wins, and little to no impact of price on the revenue-win relationship. In this case we do not need to worry about omitted variable bias. Similarly, if owners dynamically set price over the course of the season—as some teams have recently (Tishgart, 2011)—then they can respond relatively quickly to the success or failure of their team in order to maximize revenue based on demand. In this instance price will always be updated to the optimal level based on the success of the team and will not bias our estimation for the revenue generated from wins.

The second and most important testable hypothesis (H2) is that the team to team variation in the revenue-win relationship cannot fully be explained by market size, market demographics, division, or other observable characteristics. Usually many of these factors are used to control for different baseline revenues; however, it is plausible that they also affect the marginal return to a win. I will analyze how much of the variation in the revenue-win relationship can be explained by observable characteristics and how much of it is explained by other factors difficult to quantify and unique to each team. Previous studies have relied on the assumption that team to team variation in the revenue-win relationship is due to differences in market size; H2 implies not only that this is not the case, but that even accounting for other observable characteristics is not enough to explain the unique returns to winning for each team. Therefore, models based on the market size theory would be misspecified and yield biased results, with the size of the bias depending on how much of the variation is unaccounted for when only controlling for the effect of differences in market size.

The third hypothesis (H3) is that the three different collective bargaining agreements (CBAs) negotiated during the 1995-2010 period covered in this analysis had significant, unique, and team dependent effects on the return to a marginal win. In order to promote competitive balance in hopes of increasing fan interest and league wide revenue, MLB has a system of revenue sharing through which wealth is redistributed from teams with the most revenue to teams with the least. All teams' local revenue is taxed at an equal rate and this money is pooled and then redistributed (2007-2011 Major League Collective Bargaining Agreement). Additionally, there is a "luxury tax" on payrolls exceeding a certain threshold, the proceeds from which are used by MLB for general league purposes (Brown, 2010). As Burger and Walters (2003) point out, the effect of these measures is to decrease the value of a marginal win for all teams because the added revenue of a win will be taxed at a higher rate for each team than the rate at which the money will return in the form of redistribution. For teams in the middle of the revenue distribution there is the potential double effect of taking away revenue from the win and, if the added revenue is great enough, turning the club from a net payee club to a net payor club. Net payor clubs are those teams for whom revenue is great enough that revenue sharing results in a net loss. Under the first CBA in the sample, teams with above average revenues received a smaller percentage of the redistributed tax revenue (1996-2001 Major League Collective Bargaining Agreement). Under the last two CBAs, net payor clubs received less revenue from the MLB Central Fund than did net payee clubs (2007-2011 Major League Collective Bargaining Agreement; Brown, 2006). The marginal tax rate on revenue and the system of redistribution has varied across CBAs, so the amount of deflation of the value of a win has changed depending on the agreement in place, and the magnitude of the effect for teams at different places in the revenue distribution has varied depending on the agreement.[1]

III. Hypothesis Testing Strategy

The first of two models I will take to the data to test the above hypotheses is based upon the main two assumptions of team specificity and nonlinearity of the revenue-win relationship. Model 1 will allow me to test H1 and the assumption of significant variation across teams in the returns to winning. The model approximates the commonly used model for estimating the revenue-win relationship, but uses a different functional form for the relationship, includes price as a control variable, and omits other controls common to earlier models. The controls omitted are factors that are relatively stable over time and generally outside of the team's control, but thought to influence the revenue-win relationship. My

[1] The 1997-2001 "Basic Agreement" outlined a rather complicated system of redistribution and employed a 20% marginal tax rate on teams' local revenue. For the first two years of the CBA a "Hybrid Pool Plan" was in place, under which teams received the more favorable outcome of a "Split Pool Plan" or a "Straight Pool Plan". A split pool is where the tax revenue is redistributed unequally, with the higher percentage going to teams with below average revenue. The straight pool plan redistributes the revenue equally, with lower revenue teams being net payees because under the flat marginal tax rate they pay less into the pool due to lower local revenue. The split pool obviously is more generous to lower revenue teams and potentially has the unintended effect of de-incentivizing clubs near the league average from increasing revenue so that they receive a smaller percentage of the redistributed revenue (1996-2001 Major League Collective Bargaining Agreement).

Since 2002 MLB has used a straight pool plan, but has changed the marginal tax rate to 34% and 31% in the respective CBAs. Additionally, MLB Central Fund revenue generated from national broadcasts, online sites, and online media have been distributed unequally in favor of payee clubs. MLB has used a formula to derive "performance factors" for each team in order to decide the percentage of Central Fund revenue they will receive (2007-2011 Major League Collective Bargaining Agreement). Teams with the lowest revenue receive more, again de-incentivizing teams to increase their own local revenue, as this will decrease the amount of revenue they receive from the league.

estimate for the revenue-win relationship from Model 1 picks up the effect of these variables, then in Model 2 I will estimate how much of the variation in the revenue-win relationship is explained by these variables alone. Therefore, we arrive at Model 1 to estimate the revenue-win relationship:

$Revenue_{it}=\beta_0 Team_i + \beta_1 Team_i *WinPct_{it} + \beta_2 Team_i*WinPct2_{it}^{1/2} + \beta_3 Price_{it} + \beta_4 Stadium_{it}^{-1} + B_5 Stadium_{it}^{-2} + B_6 New_{it}^{-1} + B_7 New_{it}^{-2} + \beta_8 Playoffs_{it-1} + \beta_9 t + \varepsilon_{it}$, 1)

where *i* indexes team, *t* indexes year, ε is a well behaved error term and:

- $Revenue_{it}$ is the revenue for team i in year t. Revenue is defined as the team's revenue in real 2010 dollars, net of revenue sharing
- $Team_i$ is a binary variable for each of the 29 teams in the sample.
- $WinPct_{it}$ is the win percentage of team i in year t.
- $WinPct2_{it}$ is 0 if a team is below the win percentage threshold of .506 and is the added win percentage above the threshold for teams above it. The threshold was set at the lowest winning percentage of a team reaching the playoffs in the sample, the 2005 San Diego Padres.
- $Price_{it}$ is the Fan Cost Index according to Team Marketing Report for team i in year t. The Fan Cost Index "comprises the prices of four adult average-price tickets, two small draft beers, four small soft drinks, four regular-size hot dogs, parking for one car, two game programs and two least expensive, adult-size adjustable caps" (Team Marketing Report, 2011).
- $Stadium_{it}$ is the age of the stadium of team i in year t up to 6 years; any stadiums over 6 years old in year t have a value of 0 for the stadium variable, as Clapp and Hakes (2005) find the honeymoon effect for a new stadium tends to last about 6 years.
- New_{it} is the number of years team i has been in the league in year t up to 6 years, at which point the variable becomes 0. This variable captures the effect on revenue of being an expansion team and the excitement that accompanies a new team. Both the stadium variable and the new team variable are in the form of a negative polynomial to account for diminishing marginal revenue over time for these effects.

- $Playoffs_{it-1}$ is a dummy variable indicating whether or not team i reached the playoffs in year t-1.
- t is a time trend variable to capture the growth of revenue across baseball as a whole during the period studied and accounts for serial autocorrelation of the errors.

An F-test for joint equality of the 29 team estimations of $\beta_1 + \beta_2$ will test the validity of the first assumption that teams have unique returns to wins. Similarly, the size and significance of β_2 will indicate whether or not the revenue-win relationship truly is nonlinear for teams, as this would represent a significant additional increase in revenue from a win once teams are in playoff contention. The significance and magnitude of β_3 will indicate the average effect of a change in prices on revenue, but more importantly the size of the change of the estimated coefficients of β_1 and β_2 when the model is estimated with and without the price variable will indicate whether or not price is indeed an omitted variable we need to control for. A significant change in these coefficients in models that exclude this variable indicates that some of the effect of wins on revenue estimated by the coefficients is actually the effect of price on revenue through its correlation with wins.

The second model captured in equation 2 addresses H2 by estimating how much of the variation in the revenue-win relationship can be explained by factors relatively stable over time, out of the team's control, and assumed to influence the revenue-win relationship. The model is as follows:

$$Coefficient_i = \gamma_0 + \gamma_1 Market_i + \gamma_2 Income_i + \gamma_3 Division_i + \gamma_4 Distance_i + \gamma_5 Sports_i + \mu_i, \qquad 2)$$

where *i* indexes team, μ is a well behaved error term, and:

- $Coefficient_i$ is (separately) the estimates for β_0, β_1, and $\beta_1 + \beta_2$ for team i from Model 1, yielding three different specifications of Model 2. For the specification where β_1 alone is the dependent variable I will use the estimate for β_1 from a regression of Model 1 that does not include the threshold win percentage variable. Therefore, in this instance β_1 will represent each team's returns to winning assuming that the revenue-win relationship is purely linear for each team. Equation 2 will be fit with weighted least squares, giving more weight to teams for whom the estimated coefficients have smaller standard errors and therefore for whom there is more certainty about the revenue-win relationship. The three specifications will provide insight on how a team's baseline revenue (β_0), linear revenue-win relationship (β_1), and nonlinear revenue-win relationship (β_1 + β_2) are related to the independent variables in Model 2.

- $Market_i$ is the time average population of the market for team i and proxies for market size.

- $Income_i$ is the time average real 2010 per capita income for the residents of the city of team i.
- $Division_i$ will be a set of dummy variables for each of the 6 divisions within baseball, leaving out the American League Central.
- $Distance_i$ is the distance to the closest MLB team for team i.
- $Sports_i$ is the average number of other major pro sports franchises (NBA, NFL, NHL) in the market for team i during the sample.

The R^2 from this regression will indicate how much of the variation in the revenue-win relationship between teams can be explained by the above factors. Whatever variation cannot be explained is due to other factors that are so team specific they cannot be controlled for with a simple interaction term in the first model. Additionally, I will look at how the R^2 changes when I remove each of the independent variables one at a time. The change in R^2 will indicate how much of the variation in the revenue-win relationship across teams is accounted for by the independent variable that was removed from equation 2; therefore, this process will indicate which variable accounts for the greatest amount of variation across teams.

As a robustness check I will categorize teams that appear to have similar returns to winning and then interact these categories with indicators for the time period of the three different collective bargaining agreements during the sample. If the different CBAs had a significant impact on the returns to winning that varied similarly across teams with similar revenue-win relationships then we should see significantly different estimates for the wins coefficients of teams in the same categories in different time periods. This method should also indicate whether it is indeed possible to actually group similar teams or whether teams indeed have such unique relationships with their fans that they must be considered on a case by case basis. The categorized specification for Model1 is then as follows:

$$Revenue_{it}=\beta_0 Category_j*CBA_t+ \beta_1\ Category_j\ *WinPct_{it}*CBA_t + \beta_2 Category_j*WinPct2_{it}^{2}*CBA_t + \beta_3 Price_{it} + \beta_4 Stadium_{it}^{-1} + \beta_5 Stadium_{it}^{-2} + B_6 New_{it}^{-1} + B_7 New_{it}^{-2} + \beta_8 Playoffs_{it-1} + \beta_9 t + \varepsilon_{it} \quad 3)$$

- Here $Category_j$ for j=1,…, 5 is a dummy variable indicating the 5 categories of teams based on their return to a win.
- CBA_t is a three-element vector of binary variables indicating the collective bargaining agreement time periods 1997-2001, 2002-2006, or 2007-2010.

In this model there will be 3*j=15 estimates for β_0, β_1, and β_2. Each coefficient estimate has the same interpretation as equation 1 except that instead of being purely team specific and the average over time, each of the first three coefficients are now the time average for a given collective bargaining period and a given group of teams j.

IV. Data and Results

Data Sources and Coding of Specific Variables

In order to test my three hypotheses I have a panel dataset for the sixteen year period beginning in 1995, following the strike season of 1994, through the 2010 season. Each team therefore has sixteen observations for each variable in Model 1 with the exception of Tampa Bay (TB) and Arizona (ARI) which entered the league in 1998 and thus have thirteen observations. For each team in the MLB during this period—except the Montreal Expos who moved to Washington D.C. in the middle of the period and are omitted from the analysis—I use the yearly estimates for total team revenue from Forbes' valuations of MLB teams for my estimate of team revenue. All variables representing dollar values are in real 2010 U.S. dollars, adjusted using a yearly consumer price index (CPI) price deflator and the yearly Canadian to U.S. dollar exchange rate for the Toronto Blue Jays. Revenue estimates are in millions of dollars, while the price variable (proxied by the FCI) and per capita income are in U.S. dollars. Data on team wins and playoff appearances is readily available online at MLB.com and is used to construct the win percentage, win threshold, and playoff variables for Model 1. Descriptive statistics for the variables in Model 1 across teams can be found in Table 1 and by team in Table 2.

For Model 2 market size and per capita income data are gathered from various U.S. Census and Bureau of Economic Analysis sources, with a team's market considered to be the consolidated metropolitan statistical area (CMSA) within which a team operates. Market size is proxied by CMSA population and is in millions of persons within a given CMSA. Similar methodology is used with Canadian government data regarding the Toronto Blue Jays. The distance between MLB teams is in miles and only varies over the course of the sample for those teams who are closest to TB, ARI, and the Washington Nationals who all entered the league in the middle of the sample. Detroit (DET) and Milwaukee (MIL) switched divisions in 1998 when the MLB divisions were realigned with the inception of ARI and TB, with DET moving from the AL East to the AL Central and MIL moving from the AL Central to the NL Central; however, for the purposes of this analysis they are considered to be in their current divisions for the entire sample. All of the variables for Model 2 are then averaged over time to arrive at one value for each team for the regression analysis of Model 2.

The dependent variables in the three specifications of Model 2 are the estimates for each team from Model 1 and represent the time average baseline revenue (team specific intercepts), time average linear combination of the win percentage effect on revenue and square root of win percentage above the .506 threshold effect on revenue, and the time average linear win percentage effect on revenue omitting the threshold variable, respectively. These three variables will be referred to hereafter as the "intercept", "linear combination of win coefficients", and "linear win coefficient", respectively. Descriptive statistics for the variables in Model 2 can be found across all teams and by team in Tables 6 and 7, respectively.

Replication of Burger and Walters Results

The first robustness check of my methodology is to replicate the results generated by Burger and Walters. The main concern with my data sample is that it is uses Forbes' estimates for total team revenue net of revenue sharing, not on the internal measures of team local revenue released by MLB that Burger and Walters used. Local revenue is used to capture only the revenue that is affected by winning and omits revenue distributed to all teams regardless of their performance, meaning that by using total revenue I could be introducing measurement error in my dependent variable. To test this possibility I replicated Burger and Walters' model and its various specifications, with the results presented in Table 3. While I utilize win percentage in my model because of variation in the number of games played in a season due to play-in games to reach the playoffs, rainouts, or the shorter 1995 season, Burger and Walters use actual wins as a regressor. Therefore, I have normalized wins for each team in a given season by multiplying their win percentage by 162 and use this variable to construct the threshold variable, which takes a value of normalized wins-.506*162 for win totals above this threshold, and zero below it. I use the same measure of market size (population in a team's CMSA in millions of persons) as I use in Model 2, except that in accordance with Burger and Walters I divide population in half for teams sharing a market, which I will not do in Model 2.

In the Appendix I have included a copy of Burger and Walters' regression table to compare to Table 3, which is my replication of their model and its specifications. It is quite clear that my results mirror theirs extremely well[2], with very similar magnitudes and significance for each of the variables of interest—the main effects of market size, wins, and wins above the threshold and the interacted effects of market size with wins and wins above the threshold. Especially of note are the estimates for specification 2, which omits the main effects of market size and wins, which are the estimates Burger and Walters use for their analysis. Importantly, both of our regressions find that when omitting the main effects the

---

[2] The only notable difference in my replication and in their findings is the estimate of the effect of stadium age. They find an insignificant negative effect of a stadium's age on revenue—implying that as a stadium increases in age a team's revenue may decrease, or at least stays the same. My results indicate the opposite, with a significantly positive coefficient on the age variable. This is most likely due to the coding of my age variable, which takes values 0 to 6 due to convenience because of how it is coded for my models. If their variable has positive values for all teams then there will be teams in the sample with very high values for stadium age driving down the effect of this variable, whereas my variable only picks up effects for teams who constructed a new stadium around the time of my sample. Therefore, my age variable most likely picks up some of the positive returns to a new stadium within its honeymoon period and any negative returns beyond that are not taken into account due to the variable taking a value of zero. This should not significantly affect my estimates for the effects of winning and thus it can be confidently stated that there is no evidence that my dataset is significantly different from theirs in the areas of interest.

interacted effects are strong and significant, but when the main effects are included in the full specification the interacted effects are no longer significant.

Burger and Walters correctly attribute this to the multicollinearity associated with the main effects and the interactions between continuous variables, which indeed tends to result in large standard errors. The accepted method for dealing with such multicollinearity is to normalize each of the variables by centering them about their mean and fit the model including both the main effects and the interactions. This process reduces the multicollinearity between the main effects and the interactions, reduces the standard errors and changes estimation of the main effects, but should not change the overall estimation of the interacted effects. This is exactly what we find in specification 5 of Table 3. The interacted effects are exactly the same and remain insignificant; however, the estimated main effects change drastically. Market size, wins, and wins above the threshold all have significant positive main effects—which follows intuitively from their interpretation. After centering the variables we are now seeing the effect of market size for a team with *average* wins and *average* wins above the threshold. Similarly, we are seeing the effect of wins and wins above the threshold for a team with *average* market size. In the average these main effects across teams are significant, a result that will be replicated below in analysis of my own Model 1.

This implies—seemingly contrary to Burger and Walters' thesis—that across teams there are significant returns to winning and to larger market size, but that the deviation from this average main effect of winning caused by variation in market size is not significant. Essentially, market size does not significantly alter a team's return to winning compared to the average effect across teams. This will be analyzed further in this paper using Model 2.

Model 1 Price Effect Analysis

Table 5 presents my results for the regression of Model 1, which has team specific effects for baseline revenue, win percentage, and the square root of win percentage above the threshold. The first step of the analysis is to test H1, that price is an omitted variable that must be included in the model. To test this hypothesis I estimated Model 1 restricting the price effect to be zero and then included price in the model to get an unrestricted estimate for its effect. I then jointly tested each of the respective hypotheses that the team specific intercept (β_0 in the full specification), the team linear win percentage coefficient in the full model (β_1 in the full specification), and the team win threshold coefficient (β_2 in the full specification) are significantly different between the restricted and unrestricted models. In each case I found p-values of 0.000 indicating that the estimates are indeed significantly different, implying that price does have a significant effect on the estimated returns to winning—as well as a team's baseline revenue measured by the intercepts—and must be included in the model. It should be noted that when price is included in the full specification of Model 1 its effect is not allowed to vary by team and, therefore, its coefficient represents the average effect across teams of price on revenue. This coefficient in the full specification is significant and positive, indicating that on average an increase in price increases revenue. The inclusion of price, however, has an ambiguous effect on the various win coefficients depending on the team studied. For some teams including price increases the returns to winning, while for others it decreases the estimated returns. This implies that the relationship between price and the revenue generated from winning varies by team. This will be analyzed further in the discussion, but I note here that all further references to results from Model 1 are based on specifications that include the price variable.

Model 1 Estimation

In order to understand how winning affects revenue across teams, I estimated Model 1 pooling all teams, with results in Table 4. The results of the full specification indicate that controlling for the effects of being a new team, having a new stadium, making the playoffs the year before, and changing price, winning beneath the threshold may have an insignificant effect, but that winning above the threshold does indeed have significant revenue returns for the average team. The first figure in the Appendix is the estimated curve of revenue residualized for the effects of the control variables plotted against win percentage with a scatter plot of each of these revenue residuals. There is certainly an upward trend to residualized revenue as winning increases, but even across teams it is not clear that beneath the threshold this trend is significant.

The next step of the analysis is to estimate equation 1 with team specific effects for baseline revenue and winning, with results in Table 5, as mentioned above. One win can be interpreted as 1/162=.006 percentage points so each additional win below the threshold corresponds to β_1*.006 millions of dollars. To interpret each win above the threshold one must have a starting win percentage to which we add .006 because the effect is nonlinear. Therefore, one win above the threshold corresponds to β_1*.006 + β_2*((starting win percentage +.006 - .506)$^{1/2}$ – (starting win percentage - .506)$^{1/2}$) millions of dollars. The first notable finding is that many teams have insignificant estimates for baseline revenue and the win variables, potentially due in many cases to small sample size and thus large standard errors. Additionally, some teams have significant negative estimates for one of the win coefficients, indicating winning actually has negative effects on revenue, depending where on the revenue-win curve the team is. In some cases this could be due to chance and is driven by a few outliers in the data, but in some cases there may be a theoretical explanation for these results that will be discussed in the following section. In order to control for the possibility that winning has lagged effects on revenue, Model 1 was also estimated using one and two year lags of win percentage, but this strategy does not significantly alter the result of various significant negative coefficients.

The full specification of Model 1 fits the data as a whole remarkably well, with a true adjusted R^2 of .9534 when a constant is included and one team intercept is omitted. There is significant evidence that the returns to winning vary greatly across teams. A joint test of the equality of the linear combination of the wins coefficients for each team rejects the null of equality and the same is true of the test of the linear win percentage coefficient in specification 4 (which omits the threshold variable), indicating significant differences across teams. The Appendix contains figures for select teams of the estimated curve of revenue residualized from a regression on the control variables versus win percentage, and it is clear that these curves are unique to each team.

Furthermore, the magnitude of this variation has important implications. Using specification 4, which omits the threshold variable, we can easily see the magnitude of the difference in the returns to winning between teams. ARI is the only team in this specification with a significantly negative estimate, and therefore has the lowest significant estimate for the revenue generated by wins, with a coefficient of -79.76. BAL has the highest significant positive estimate at 321.07. This difference implies that an additional win for ARI reduces revenue by about $480,000, while an additional win for BAL increases revenue by about $1.92 million, a difference of about $2.4 million. Even if we ignore ARI's negative coefficient, there are ample teams with insignificant returns to winning, implying that a win is $1.92 million more valuable to BAL than it is to teams with zero returns to winning.

Looking at the full specification we can see the differences in magnitude based on where a team is on the win curve and which team is being considered. For example, MIL has a significant linear win percentage coefficient of 116.90 and significant win threshold coefficient of 79.50 in the full specification. This indicates that a win that takes MIL from 79 to 80 wins is worth 116.90*.006 = $700,000; however, a win that takes MIL from 89 to 90 wins is worth 116.90*.006 + 79.50*.014 = $1.84 million. This is in contrast to a team with insignificant returns to winning, which would gain zero revenue from either of these additional wins.

It should be noted that these dollar figure estimates are substantially lower across the board than the estimates usually cited for the value of a win, which tend to be around $4 to $5 million depending on the year studied. These estimates are in the ballpark, however, of those obtained by Burger and Walters whose highest estimated return to winning was for NYY at about $3.62 million per win and whose lowest was $590,000 per win for MIL. Additionally, we must consider that the goal of these estimates is to put a dollar figure on wins within the same season. As has been noted, there is evidence that wins could generate revenue the following year, but this additional revenue is very difficult to measure and would not be included in my estimates.

Robustness Checks

To validate the robustness of Model 1's finding that there is significant team variation in the returns to winning and to analyze the potential effect of the various CBAs, I estimated the categorized version of Model 1 discussed in Section 3 using two different methods to choose the categories. The assumption is that one factor that could be driving the estimates in Model 1 is that the effect of winning on revenue varies within the sample by time period based on the various CBAs. Additionally, if it is the case that there is team variation in the returns to winning but that teams can be grouped with other teams with similar characteristics, then these time period effects should be similar for similar groups of teams. The categorized version of Model 1 should then lead to stronger and more significant estimates for the returns to winning because I am clustering similar teams and allowing the effects to vary by time period.

The first method was to group teams based on the data. Plots of the estimated linear combination of the wins coefficients from Model 1 versus the estimated team intercepts from Model 1 and the estimated linear win coefficient of Model 1 versus the team intercepts both indicate a strong negative correlation between the returns to winning and a team's baseline revenue. Theoretical justification for this finding can be found in the following section and has interesting implications for analysis of the revenue-win relationship. Additionally, this finding allowed me to group teams based on their estimated returns to winning from Model 1 with the understanding that these groupings effectively clustered teams with similar intercepts as well. Therefore, in the first method there are five categories of teams beginning with a category for teams with high estimated returns to winning (and thus low estimated intercepts) and ending with a category for teams with low estimated returns to winning (and high intercepts). The second method was to group teams theoretically based on my knowledge of MLB and teams' historic relationships with their fan base. I grouped teams I assumed have similar fan bases who respond similarly to winning, without reference to the estimates from Model 1. For details of which teams were grouped together refer to the note under Table 12.

Tables 11 and 12 present the results from the two categorized specifications of Model 1, with the categories based on the data in Table 11 and the theoretically chosen categories in Table 12. The important finding is that in both cases I find very few significant estimates for the win coefficients, indicating that teams within categories do not actually have similar returns to winning, even within the three time periods. The implication is that even within small time periods teams appear to have very unique revenue returns to winning and that in many cases the returns may in fact be insignificant. This is an important finding for the analysis that will follow because it suggests that teams that on the surface appear to be similar or that have similar estimates from Model 1 cannot necessarily be grouped together. The implication is that even two teams with almost the exact same demographics and even the same time average returns to winning could have very different returns to winning in a smaller time period due to factors that cannot be accounted for in a model. Furthermore, these results suggest no real effect of the CBAs on the returns to winning within categories, because for any given category the estimates for each time period are not significantly different from zero and thus not significantly different from each other.

Model 2 Estimation

The results from Model 1 and its subsequent robustness checks indicate strong variation among all 29 teams in the manner in which winning drives revenue. Additionally, the results validate H1—that price is an omitted variable that must be included—and refute H3—that the different CBAs had significant effects on the returns to winning across groups of teams. In order to better understand where the team variation in revenue comes from, I regressed the estimated team intercepts, team linear wins coefficients, and team linear combinations of the wins coefficients on a group of explanatory variables believed to explain this variation. Of most importance is the market size variable, proxied by population, which is thought to explain most of this variation. I want to weight each of these dependent variables by my confidence in their robustness and therefore utilized weighted least squares for this analysis, with the inverse of the standard errors as weights.[3]

---

[3] Weighted least squares is a commonly accepted econometric method for regression analysis when the dependent variable is estimated rather than observed. In this case the dependent variable in the three different versions of Model 2 represent each team's estimated time average for the variable considered—baseline revenue for the intercept and the returns to winning for the other two versions. The measures for these dependent variables are the coefficients from the regression of Model 1 for each team, which have varying levels of certainty due to

Tables 8, 9, and 10 contain the regression results for Model 2 with the team intercepts, team linear combination of the wins coefficients, and team linear wins coefficients as the dependent variable, respectively. The full specification contains all of the explanatory variables and has R^2's of .4522, .5525, and .6179 for the three regressions, indicating the regressors as a whole explain about 45%, 55%, and 62% of the variation in the respective dependent variables. This implies that about 45% or 38% of the team specific variation in the returns to winning is left unexplained, depending on the specification of the effects of winning. The coefficients represent the effect of a change in each explanatory variable on a team's coefficient for the corresponding dependent variable in Model 1, holding all of the other explanatory variables constant. In essence these coefficients represent the direction and magnitude of the change in the baseline revenue or revenue generated by wins from an associated increase in population, an additional sports franchise, etc. For the division dummies the coefficients represent the difference in

---

differing levels of variation in the data for each team. Therefore, the dependent variable for each version of Model 2 has different variance for each team—i.e. the model is heteroskedastic—but this variance for each team has been estimated, so we know the form of the heteroskedasticity.

Weighted least squares (WLS) is more efficient than OLS when the form of hetorskedasticity is known. In the case where the dependent variable is estimated, WLS weights each observation of the dependent and independent variables by the inverse of the standard error of the associated dependent variable observation. In essence this method places more weight on teams that have smaller standard errors and hence more certainty about their measured dependent variable.

WLS is a special case of generalized least squares (GLS) in which the variance-covariance matrix that transforms the equation, which I will call **W** is a diagonal matrix with nonzero diagonal entries corresponding to the estimated variance of each observation for the dependent variable. Letting $\mathbf{W}^{-1}=\mathbf{V}'\mathbf{V}$, where **V** is a diagonal matrix of the inverse of the standard errors of each observation for the dependent variable, WLS transforms the regression equation $\mathbf{y}=\beta\mathbf{X}+\varepsilon$, into $\mathbf{Vy}=\beta\mathbf{VX}+\mathbf{V}\varepsilon$. This amounts to a regression of **Vy** on **VX,** and the corresponding estimator is thus a transformation from the normal OLS estimator $\hat{\beta}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ into the WLS estimator $\hat{\beta}_{WLS}=((\mathbf{VX})'(\mathbf{VX}))'(\mathbf{VX})'\mathbf{Vy}=(\mathbf{X}'\mathbf{V}'\mathbf{VX})'\mathbf{X}'\mathbf{V}'\mathbf{V}\mathrm{y}=(\mathbf{X}'\mathbf{W}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^{-1}\mathbf{y}$ (Stock and Watson, 2007). $\hat{\beta}_{WLS}$ is normally distributed, consistent—with mean equal to the true parameter β—and is more efficient than the normal OLS estimator because its variance takes the form $\mathrm{Var}(\hat{\beta}_{WLS})=(\mathbf{X}'\mathbf{W}^{-1}\mathbf{X})^{-1}$ and is thus "scaled down" from the variance of the OLS estimator by the inverse of **W**.

The Stata command vwls stands for "variance weighted least squares" and carries out weighted least squares estimation as described above when used with the option sd(var) where "var" is a variable containing the estimated standard errors for each observation of the dependent variable. The vwls command does not produce an R^2 value, but instead a chi-square goodness of fit statistic. In order to carry out the analysis needed for Model 2 I replicated the regressions using the regress command with analytic weights, specifying the inverse of the *variances* of each observation of the dependent variable as the weights. This is because the analytic weights option automatically weights the regression equation by the square root of the weights specified. This option yields the same coefficient estimates as vwls and an R^2 value that represents the amount of variation explained by the regressors, but has incorrect standard errors for the estimates. The reason is that the analytic weights option introduces a common scaling factor equal to the inverse of the mean of the weights when calculating the variance of the residuals and thus the variance of the estimates (Gould, Stata Support).

baseline revenue or the returns to winning between a team in the given division and a team in the AL Central, holding all other variables constant.

Market size has an insignificant effect on the returns to winning in both Table 9 and Table 10, indicating that, when controlling for the other variables, market size is irrelevant. Having more pro sports teams has a significantly negative effect on the returns to winning and due to the high correlation between this variable and market size, when this variable is omitted market size picks up this negative effect. We see a similar correlation for the distance variable. Teams close to another MLB team have higher returns to winning, suggesting that competition between two teams for the same fans increases the value of wins. When this effect is left out the market size effect increases and is significantly positive due to the negative correlation between market size and the distance between teams, as there are two teams in most of the largest markets. Per capita income has a significantly positive effect, meaning teams in more affluent markets have a higher return to winning. It appears that being in the NL East, a division dominated until recently by ATL, has higher returns to winning than the AL Central, a historically competitively balanced division. The opposite is true of the NL Central and NL West, which contain some teams that have not been consistently successful for an extended period of time (PIT, CIN, CHC, COL, SD, and LAD).

Each of the additional regression specifications in Tables 8, 9, and 10 leaves out one of the regressors, meaning the subsequent change in R^2 represents the amount of variation in the dependent variable explained by the corresponding omitted regressor. Most notably, we see that market size explains almost none of the variation in the win effects, as the R^2 in each of these models decreases very little when market size is omitted. Distance, number of pro sports teams, and per capita income all explain a similar amount (about 10-15%) of the variation and division explains the most, probably in part because in a small sample it is allowed to vary the most by team due to the five different dummy variables included in the regression. The overall implication is that the results refute the notion that market size is a determining factor in teams' return to winning and overall these results paired with the variation across teams in the estimates from Model 1 indicate that it is very difficult to explain this variation without looking at unique factors on a case by case basis.

V. Discussion

The overwhelming result of this analysis is that teams have a unique relationship with their fans that cannot be easily explained by their time invariant characteristics. For the average team there is a significant return to winning, but it is not clear that winning matters until a team nears playoff contention. When this effect is analyzed on a team by team basis, however, the results are much less clear. For some teams, such as PHI, winning has the assumed effect of increasing fan responsiveness and generating revenue. For other teams—see ARI—winning appears to have the opposite effect, decreasing revenue. This result is due to a number of complex factors determining whether or not fans invest monetarily in their local team, a few of which will be analyzed in this discussion to make the case that a one size fits all model for team revenue provides biased estimates for the revenue generated by winning.

The Price Effect

Price is an important factor in many fans' decision of whether or not to invest in a team, but averaged across all of the fans in a market this effect varies between teams due to the demographics of the market. A team in a wealthy market likely has enough affluent fans that increases in price are not much of a deterrent, meaning the decision of whether or not to buy tickets is based more upon the performance of the team. A winning team is more fun to watch and so paying a little extra for this added utility is a minor cost. On the other hand, this added cost may be a significant deterrent for less affluent fans. A higher fan price elasticity of demand is the most likely explanation for the negative linear win percentage coefficient in specification 4 for Arizona. Phoenix is the least affluent market in the U.S. in the sample, implying that price may be a significant factor in the average ARI fan's decision of whether or not to attend games. Across teams the price variable has a positive coefficient, so an increase in price should increase revenue (holding all else constant), but in ARI a price increase may well deter many former consumers from buying tickets, even in a winning year. Therefore, if it is the case that ARI raises prices in years in which it expects to be successful, then this price increase offsets the added revenue that winning should generate, by decreasing demand and attendance. If price is indeed correlated with ARI's actual performance, then in these years price will increase and wins will increase, but revenue will decrease or stay about the same depending on the relative magnitudes of the price change and decreased demand. Given that the model estimates an increase in revenue from the price increase, ARI's win coefficients are driven downward to offset this increase in order to minimize the difference between predicted revenue and actual revenue. In effect, the win coefficients for ARI really represent its fans' deviation downward from the league average price effect, instead of actually picking up the returns to winning. In this way the specification of the model does not fit ARI well and leads to an incorrect estimate for the team's return to winning, but this also signifies the complexity of analyzing the factors that determine the value of a win because this is an issue that teams in wealthier markets do not face.

Perpetual Winning or Losing

In *Diamond Dollars* Vince Gennaro refers to the diminishing returns to making the playoffs many years in a row and specifically cites ATL as an example of this effect. Similarly, one can conjure examples of cities once thought to be great baseball markets now considered business flops after many successive years of losing—see PIT and BAL. My estimates from Model 1 support the theory that repetitive winning or losing diminishes the returns to winning as fans grow accustomed to the same performance from their team year in and year out.

On one end of the spectrum we have ATL, which plays in a relatively large market, has a consistent following across the south because for a long time it was the only team south of the Mason-Dixon line, and which should theoretically be able to fill the stadium at Turner Field when they have a high quality product. Unfortunately—from a business perspective—ATL won their division and made the playoffs every year from 1991 through 2005. The result is that by the time we reach my sample, fans in ATL had grown accustomed to the Braves winning, going to the playoffs, and inevitably failing to win the World Series, as they only won one world championship during this span. This lack of responsiveness to winning is seen in ATL's estimates, which are negative above the threshold. The negative coefficient may be due mostly to the specification of a nonlinear curve, but it appears from their figure in the Appendix that, at most, wins above the threshold generate zero revenue.

Conversely there is PIT, a team with a storied franchise history, known for their prowess in the 70's and early 90's, in addition to Bill Mazeroski's 1960 World Series clinching home run against the Yankees. During my sample PIT is the only team to fail to reach the win threshold and is regarded as a franchise notorious for trading away players at their peak rather than ponying up higher salary costs to keep them. The construction of a new stadium in 2001 was heralded as a fresh start for the franchise, but the ten years of losing that followed ensured that fans in Pittsburgh are unwilling to watch a failing team, even if they can do so while enjoying modern amenities. This phenomenon is evident in PIT's negative win coefficient, which represents that fans have grown tired of watching bad teams, leading to insignificant returns from winning.

Specification of the Model

Given the small sample sizes for each team, teams with few data points on either side of the threshold are subject to the bias presented by a few outlying years. For example, OAK has a strong negative linear coefficient in the full specification, but looking at their figure in the Appendix it appears that this is due to an unusually high level of residualized revenue in their least winning year, 1997. Overall, OAK appears to have an insignificant return to winning, which is supported by their insignificant estimate for the linear coefficient in specification 4 when the threshold term is omitted. Therefore, the specification of the model when forced upon OAK's sixteen years of data is skewed by this one year. We see a similarly high outlying data point for CWS in 2007, despite the team's lowest win percentage during the sample. In this case there is a natural explanation for the outlier: two years earlier in 2005 the team won its first World Series since 1917 and the associated excitement and accompanying boost in revenue carried over not only to 2006, but to 2007.

Capacity Ceiling

One of the more interesting developments of the analysis is the negative linear correlation between a team's baseline revenue and its returns to winning. The implication of this finding is that as a team's baseline level of revenue increases, winning generates less revenue. This may seem intuitive, but it does not have to be the case and indeed many other studies would suggest the opposite. If it were true that each market had the same proportion of bandwagon and purist fans and thus large markets had more of both, then teams in large markets should have both the highest baseline revenue *and* the highest returns to winning. This appears not to be the case, however, as my results suggest the opposite relationship between baseline revenue and revenue generated by wins.

There are two potential reasons for this result, depending upon the team studied. The first, and most important, is that there exists a natural capacity constraint on a team's revenue due to the limited number of seats in a stadium. For BOS and NYY—who are consistently playoff contenders throughout the sample and have notoriously large, rabid fan bases—it may be that there is simply no more room for them to significantly increase revenue year to year. In some ways this should not be the case, as ownership of their own regional sports network (RSN) means that, ideally, increased winning should lead to more demand for their television product and should increase revenue. On the other hand, given that each year they are in contention, fans may tune in regardless and so the capacity constraint at the ballpark limits the amount of revenue they can make from ticket sales, concessions, etc. and the revenue generated from television advertising remains stable. The result is seemingly insignificant effects on revenue from winning and large baseline revenue that is constant from year to year.

OAK is an example of a team with an unusually high intercept despite having in actuality low baseline revenue, rarely selling out, and thus having no true capacity constraint to speak of. The reason for this result is simply the specification of the model. As previously mentioned OAK has a strong, negative linear wins coefficient due mostly to one outlying data point. In reality OAK probably has a fan base that is not very responsive to baseball due to other substitutes in the area, an old worn down stadium, and low disposable income. The result is a stable, relatively low amount of revenue generated each year. The model's estimate for the amount of predicted revenue for OAK is driven downwards each year, however, by the negative wins coefficient. To offset this effect a high intercept is needed so that once this negative trend is accounted for the predicted revenue matches the actual revenue. Indeed this is exactly what we see: the intercept for OAK in specification 4, which leaves out the threshold term, is significantly lower than the one in the full specification; the wins coefficient in specification 4 is also higher than the one in the full specification, meaning that OAK moves from a high intercept, low returns to winning team in the full specification to an average intercept, average returns to winning team in specification 4. Therefore, when the model is aligned more closely to OAK's actual situation we see an estimated intercept and estimated returns to winning that match this reality.

For the teams in the sample which do not sellout every year we may be seeing varying degrees of the capacity constraint effect. Teams that are consistently close to selling out have relatively large baseline revenues, but still have some room to increase revenue when they win. The amount of revenue generated from winning may be capped, however, once they start selling out regularly. Teams with smaller followings will have smaller baseline revenues and more room to increase revenue as they win, thus allowing for larger returns to winning. While this effect likely drives the general relationship between the intercept and the win coefficients, it is not the only factor determining a team's return to winning.

Insignificance of Market Size

The goal of this paper was to understand more fully the characteristics of a team that determine their fans' responsiveness to winning, based on the underlying theory that market size would be the most important of these characteristics. All of the evidence presented so far consistently rejects this theory. Market size not only appears to have an insignificant effect on a team's estimated returns to winning, it explains almost none of the variation between teams in these estimates. Lest one feel that this is due to somehow biased estimates, I will analyze this phenomenon further anecdotally by comparing the teams that share two of the largest markets, Los Angeles and Chicago.

In Los Angeles we find two very different looking franchises. The Angels, technically located in Anaheim, have been trying to assert themselves in the LA market for a long time, even changing their name from the Anaheim Angels to the Los Angeles Angels of Anaheim. Winning seems to have a profound effect on their fans, which is in part why the team recently signed Albert Pujols—known as the best hitter in the game—to a lucrative long term contract. On the other hand, their counterparts the Dodgers appear to have very low, if any, returns to winning. Whether it is a capacity constraint issue, the tastes of their fans given the other entertainment available in Southern California, or the team's prolonged lack of success in the playoffs, the team appears to have relatively constant revenue regardless of performance.

In Chicago there are two teams with again very different business success and unclear returns to winning. The Cubs are perpetually near the top of baseball in terms of value and revenue despite limited success. The full specification yields a significant negative coefficient beneath the threshold and a significant positive one above it, but leaving out the threshold term yields an insignificant linear coefficient on win percentage. The team's plot seems to support the conclusion of limited returns to winning, probably in part due to a capacity constraint similar to that faced by NYY. On the other hand, CWS appears to have insignificant returns to winning by any measure, but also makes significantly less money than the Cubs each year.

By comparing teams within large markets it is clear that a large market does not significantly increase the returns to winning. The results of Model 2, however, suggest that increasing the distance between teams diminishes the returns to winning, implying that having a team very close by should increase the returns to winning. The underlying implication is that the impact of sharing a large market with another team is unique to the team being studied. In each of the two markets discussed above we see a more established team making revenue each year regardless of performance and one less established team fighting to gain support; in the case of LAA this support is based in large part on performance. It is certainly not the case that teams within the same market will make the same amount of revenue or will have the same returns to winning, however. This analysis supports what the rest of the paper has argued—that a team's relationship with fans is extremely unique and independent of market size.

VI. Caveats

Potential Unobserved Factors

Correctly estimating the returns to wins in one season for a team is clearly very difficult due to the complexity of a number of underlying factors. The goal of the discussion so far has been to explore the relevance of some of these factors with respect to their roles in the model. For the majority of teams, however, it appears that there is a relatively insignificant effect of winning on revenue or that this effect cannot be weeded out from all of the other noise. I will provide three possible further explanations for this "noise" in an effort to provide insight into the factors that teams would need to consider on a year to year basis to have a more complete understanding of their own revenue generating process.

The first factor is the lagged effect of winning. I attempted to control in part for this in Model 1 by including lagged win percentage variables, but in reality this effect—like the effect of current wins—probably varies team by team and year to year. As we see with CWS, winning the World Series is a big deal to fans—so big that it probably has revenue effects many years into the future. On a smaller scale making the playoffs and even having a winning season can have future effects that stretch beyond even the next season. What is difficult to estimate in a model is not only what winning a championship means in terms of fan responsiveness, but how it affects the business as a whole. Owners are likely to use a perceived bump in fan demand to establish other initiatives to increase this momentum, such as stadium renovations, more promotional deals, or increased activity in the free agent market. In the case of signing free agents, these moves may not only increase revenue, but could also generate wins many years in the future. It would be almost impossible to partial out the effect of past wins from corresponding business improvements in a model fit across all teams.

A similar factor is the revenue generated by television deals. Some teams with large enough fan bases own part of a RSN, which has been shown to generate a windfall of revenue year to year for teams with an established station. In fact this status was considered as a variable to include in Model 2 in order to understand just how RSN ownership may affect the returns to wins; however, there are ample cases of teams creating RSNs only to have them fold after a few years—see KC. Therefore, in many ways this would be an endogenous variable, with the returns to winning potentially determining whether a team owns a RSN, rather than vice versa.

For the majority of teams, which do not own RSNs, television revenue comes from deals signed many years in advance with a station's local affiliate. Winning certainly impacts this revenue, as a better on field product provides more leverage in these negotiations, but given that these contracts are often for many years at a time it is almost impossible to put a dollar figure on any one win's contribution to this revenue. Indeed LAA's signing of Albert Pujols is a great example of this phenomenon. Many analysts support the notion that the high price the team paid to secure his services well beyond his peak performance years was because the team's television deal is up for negotiation soon. Owner Arte Moreno is hoping that a successful team with a superstar player will be enough to attract a large audience and give him ample leverage to drive up the value of this television contract.

The final factor is the unique historical relationship and reputation a franchise has with its fans. Baseball fans are not created equal and it is clear from this analysis that each market is not made up of the same types of fans. STL is often noted as having the "best" fans, a reference to the fact that fans seem to show up independent of the result on the field—a theory supported by my estimates and their plot in the appendix. TB is known to have fans that rarely show up regardless of result—which is supported somewhat by my results, although they have a very small sample of years above the win threshold. This relationship, however, is prone to change based on repeated performance—exemplified by PIT. Similarly, it may be possible to create a certain type of fan base over the course of many years. A Yankee executive with whom I talked mentioned that there was not a strong Yankee fan base in the 80's and early 90's after years of on field disasters. Then owner George Steinbrenner believed, however, that if he kept putting enough money into the team and demonstrated an interest in winning, eventually the fans would respond with the level of fervor generated by the great teams of the 50's and 60's (Cashman, Interview). After winning five World Series during my sample, the Yankees are known for the largest fan base in MLB, are consistently the highest grossing and most valuable team in the league, and are referred to by competitors as the "Evil Empire" for their ability to make and spend money.

Small Sample Size and Specification of the Model

The reason the aforementioned unobserved factors are important to any discussion of my analysis is that for the team by team analysis, the sample size is small, with only sixteen observations for Model 1 and twenty nine observations for Model 2. As shown previously, this means that in many cases estimates are driven by just a few data points and that there is substantial uncertainty in these estimates, resulting in large standard errors and inconclusive findings for the significance of the causal relationships studied. In Section 5 I discussed a few examples of teams with significant estimates and possible explanations for the observed results—either factors that could be confounding the estimates, or reasons why the estimates are theoretically justifiable. It could be, however, that in the majority of cases these estimates are just a result of the noise created by year to year variation in unobserved factors and because of the small sample size my estimates are essentially making something out of nothing. This problem is magnified by the full specification of Model 1 because the inclusion of the win threshold essentially fits two different relationships between wins and revenue—the one above the threshold and the one below it. Additionally, I am forcing wins above the threshold to have a diminishing effect for all teams by taking the square root of the variable. In essence I am taking an already small sample size, breaking it in half, then forcing a certain type of curve to fit it. In some cases it could be that this curve is not correctly specified, yielding biased results. More importantly, given the number of teams in the sample and the relatively small number who have significant estimates, it could be that these "significant" results are merely chance occurrences in which insignificant noise happened to fit the specified curve in this instance.

The issue of small sample size and the noise associated with many of my estimates is a legitimate concern for the analysis provided in this paper. There are, however, many silver linings to this caveat. The first is that the dataset as a whole is the best possible representation of the data needed for this analysis. While it would be great to have local revenue figures released by the teams themselves, this is not possible. It has been demonstrated through the very close replication of Burger and Walters' findings and various robustness checks that the measures used for revenue and wins in this analysis are solid. Furthermore, it would be unhelpful to increase the sample size by including data for years before 1995 for a few reasons. The first reason is that Forbes valuations only stretch back to 1990, so it would only be possible to add five years of data without changing the source for the data and introducing possible measurement error. More importantly, the strike of 1994 drastically changed the business of Major League Baseball and attempting to control for this effect in order to increase the sample size by five years would likely introduce more problems than it would solve.

Given that the sample used is legitimate, the actual results suggest that the analysis provided is useful, even with the small sample size. First, the pooled model used to replicate Burger and Walters and the one used as a robustness check for Model 1 both provide results consistent with my findings and have no sample size issues. The replication of Burger and Walters' model suggests that there is a significant interacted effect between wins and market size, but this effect goes away when the main effects of these variables are included. This is consistent with what I find in the rest of my analysis. The returns to winning vary significantly across teams and thus deviate from the average effect of winning on revenue. This deviation, however, is not explained by market size. Additionally, the pooled version of Model 1 suggests that wins do increase revenue on average, but this average effect may not even be significant beneath the threshold. This indicates that for many teams the effect of wins on revenue as specified in my model is insignificant, which is exactly what we find in the team specific version of Model 1.

Furthermore, specification 4 of Model 1—in which the threshold variable is left out—provides some clarity to the results. By removing the threshold term the model is left to fit a simple linear relationship for each team for all of the data in the sample. The result remains similar: for many teams the findings are inconclusive, with an insignificant coefficient; for some teams we see a significant positive relationship, as expected; and for ARI the significant negative relationship persists, but an explanation for this anomaly has already been discussed. The results for Model 2 remain the same when these linear coefficients are used as the dependent variable, indicating that doing so does not change the source of the variation across teams.

If it is the case that my estimates for Model 1 are in some way biased or unreliable due to small sample size and the specification of the model, then much of the analysis that follows in Model 2 is of little use. The overall arguments of the paper, however, remain intact regardless of the validity of some of these estimates. Across teams there is a complicated, but positive relationship between wins and revenue. When this relationship is analyzed team specifically, however, this is not the case for all teams. On a team by team basis this relationship is, at worst, driven by unobserved factors leading to a lot of noise and inconclusive findings. At best, it is represented by my estimates, which indicate a unique relationship for each team. In either case the point is essentially the same—one cannot provide a single dollar figure in any given year for the value of a win to a MLB franchise. The true value of this win depends on the team, the number of wins the team will have otherwise, and probably on a host of other year specific factors. Additionally, the value of this win is not driven by the size of the market of the team studied, and in fact is mostly driven by a unique relationship a team has with its fans that is based on factors that are dynamic in nature and impossible to control for across teams.

VII. Conclusion and Further Research Opportunities

The goal of this paper was to better understand the relationship between winning and revenue for MLB franchises in an attempt to critique current and former claims about the value of wins and, subsequently, the value of players. The underpinning hypothesis of the paper was that this revenue-win relationship is too unique to each team to make blanket statements about a win's value and that this variation between teams would persist when controlling for demographic factors. The resulting analysis not only supports this hypothesis, but demonstrates just how complicated and unique this relationship is. I find significant variation across teams in the returns to winning, variation that persists through a variety of specifications. This variation is not due to market size—in fact market size explains essentially none of the difference in estimates across teams—and is only somewhat explained by other demographic factors. Further analysis of the results for teams that share markets indicates that not only are different markets made up of fans with different responsiveness to winning, but teams within the same market have different types of fans. In most large cities with two teams it appears that there is an established team with a high revenue base that varies little with performance and there is a less established team for whom revenue varies much more dramatically.

The analysis provided in this paper is based on the assumption that owners are rational, profit-maximizing agents and therefore would use information about the revenue generated from wins when making pricing and free agent decisions. This assumption implies that in a given season if an owner could measure a player's added wins based on performance and could measure the revenue generated from these wins, then the owner would be willing to pay the player a salary up to that revenue figure. There may be another factor not considered in this analysis that provides an interesting opportunity for further research. One MLB executive mentioned that owners do not only think about maximizing profit year to year, but instead consider long run implications when making decisions. During my sample the value of MLB teams has increased each year at a rate that far outpaces inflation. Just recently the Dodgers were sold to an ownership group for a record $2 billion despite an old stadium and having failed to reach the World Series in over twenty years. The current trend indicates that given the scarcity of MLB teams available for sale and the apparent intrinsic value associated with owning a sports franchise, individuals or groups with enough money to buy a team for sale can turn a handsome profit on the investment upon sale—even with little actual value added. This profit increases as the revenue potential and prestige of the franchise increase, providing incentives for owners to bolster the competitive viability of the franchise. Therefore, in many cases, owners may pursue opportunities that create a short run loss if they are perceived to increase the long run sale value of the team (Cashman, Interview).

A great example of this is, again, the Albert Pujols signing. If Pujols plays as expected during the life of his contract he may provide a surplus of revenue in the early years while still in his prime, but will surely be a net loss in terms of revenue generated by his added wins during the final years. Whether this difference evens out could be debated, but even if Pujols' performance value is far exceeded by his salary the deal may ultimately be profitable for the team. Having Pujols on the team increases the prestige and interest in the Angels. Furthermore, the team is poised to contend for the next few years and adding Pujols greatly increases their chances of multiple playoff runs and opportunities to win a World Series. All of these factors will likely drive up revenue—especially considering their positive estimated returns to winning from my analysis—and increase the team's leverage in television negotiations. Even if this increased revenue still does not make up for Pujols' salary, however, when owner Arte Moreno decides to sell the team he will undoubtedly make up the difference, and then some. Typically MLB teams sell for about two to two-and-a-half times the amount of revenue they make (Cashman, Interview; Duquette, Interview). The implication of this is that a move such as signing Albert Pujols not only increases revenue in the short run, but if this increase persists into the future the effects will by amplified should the owner choose to sell the team. Assuming that owners consider such analysis, further research would attempt to analyze the relationship between winning and a franchise's overall value in order to compare this relationship to that between winning and short term revenue generated.

Appendix

Table 1: Summary Statistics Across Teams for Stage 1 Regression Variables

Variable	Mean	Number of Observations	SE	Minimum	Maximum
Real 2010 Revenue (Millions of Dollars)	153.06	458	57.13	35.63	448.23
Win Percentage	0.50	458	0.07	0.27	0.72
Square Root of Win Percentage Above Threshold*	0.11	458	0.13	0.00	0.46
Dummy Variable for Reaching the Playoffs in Previous Season	0.26	458	0.44	0.00	1.00
Inverse of the Number of Years Since Team's Inception**	0.01	458	0.09	0.00	1.00
Inverse of the Number of Years Since Opening a New Stadium**	0.09	458	0.21	0.00	1.00
Fan Cost Index (Real 2010 Dollars)	174.05	458	40.45	112.06	417.62

*Win percentage threshold is set at .506.
**Number of years is capped at 6, so after 6 years in existence or 6 years with a new stadium variables take value of 0.

SOURCE:
Revenue data from Forbes yearly valuations of MLB franchises. Revenue is converted to real 2010 U.S. Dollars using consumer price index (CPI) price deflator.
Win percentage, playoff, and new team data from MLB.com.
Stadium data from ballparksofbaseball.com.
Fan Cost Index is used to proxy price and is released yearly by Team Marketing Report to track the cost of attendance at one game for a family of four for each MLB team.

Table 2: Time Average of Stage 1 Regression Variables by Team

Team	Real 2010 Revenue (Millions of Dollars)	Win Percentage	Square Root of Win Percentage Above Threshold*	Dummy Variable for Reaching the Playoffs in Previous Season	Inverse of the Number of Years Since Team's Inception**	Inverse of the Number of Years Since Opening a New Stadium**	Fan Cost Index (Real 2010 Dollars)
ARI	160.504	0.491	0.106	0.308	0.188	0.188	149.672
ATL	180.233	0.575	0.245	0.688	0.000	0.153	173.179
BAL	165.708	0.458	0.032	0.125	0.000	0.039	171.144
BOS	202.183	0.561	0.220	0.563	0.000	0.000	262.259
CHC	175.959	0.490	0.087	0.250	0.000	0.000	215.196
CIN	120.957	0.469	0.060	0.063	0.000	0.153	144.710
CLE	168.838	0.528	0.154	0.438	0.000	0.091	175.019
COL	158.241	0.483	0.054	0.188	0.059	0.153	161.831
CWS	145.290	0.518	0.105	0.188	0.000	0.023	189.483
DET	135.124	0.441	0.039	0.063	0.000	0.153	174.217
FLA	111.327	0.485	0.059	0.125	0.059	0.000	145.153
HOU	152.680	0.524	0.131	0.375	0.000	0.153	183.986
KC	107.430	0.423	0.005	0.000	0.000	0.000	138.701
LAA	150.152	0.531	0.149	0.375	0.000	0.000	143.420
LAD	185.514	0.528	0.140	0.375	0.000	0.000	179.153
MIL	119.012	0.462	0.019	0.063	0.000	0.153	140.343
MIN	109.883	0.501	0.118	0.313	0.000	0.063	152.484
NYM	192.483	0.508	0.109	0.188	0.000	0.094	207.035
NYY	287.941	0.598	0.297	0.875	0.000	0.094	232.270
OAK	118.701	0.521	0.132	0.313	0.000	0.000	158.313
PHI	146.448	0.506	0.114	0.188	0.000	0.153	179.429
PIT	111.801	0.426	0.000	0.000	0.000	0.153	147.700
SD	128.704	0.492	0.084	0.250	0.000	0.153	159.933
SEA	167.292	0.508	0.132	0.250	0.000	0.153	188.410
SF	160.218	0.523	0.151	0.250	0.000	0.143	193.914
STL	157.092	0.536	0.164	0.500	0.000	0.143	186.074
TB	131.206	0.439	0.055	0.077	0.188	0.000	159.722
TEX	157.901	0.501	0.087	0.188	0.000	0.091	161.535
TOR	127.149	0.492	0.068	0.000	0.000	0.010	165.975
Total	152.965	0.501	0.107	0.261	0.017	0.086	173.802

*Win percentage threshold is set at .506.
**Number of years is capped at 6, so after 6 years in existence or 6 years with a new stadium variables take value of 0.

SOURCE:
Revenue data from Forbes' yearly valuations of MLB franchises.
Win percentage, playoff, and new team data from MLB.com.
Stadium data from ballparksofbaseball.com.
Fan Cost Index is used to proxy price and is released yearly by Team Marketing Report to track the cost of attendance at one game for a family of four for each MLB team.
Revenue and FCI are converted to real 2010 U.S. Dollars using consumer price index (CPI) price deflator.

Table 3: Burger and Walters Replicated Regression Results

	Team Revenue				
Dependent Variable	1	2	3	4	5
Adj. R-Square	0.78	0.78	0.78	0.77	0.77
Number of Observations	458	458	458	458	458
Constant	41.99**	41.79**	35.92**	37.16	85.58**
	(3.80)	(3.73)	(12.71)	(30.69)	(2.44)
Time Trend	7.93**	7.93**	7.93**	7.93**	7.93**
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)
Market Size	-0.9			-0.3	8.94**
	(3.12)			(7.33)	(0.60)
Wins			0.1	0.08	0.55**
			(0.18)	(0.43)	(0.19)
Wins Above Threshold			-0.49	-0.47	1.22**
			(0.59)	(0.87)	(0.38)
Market Size*Wins	0.11*	0.09**	0.09**	0.09	0.09
	(0.04)	(0.01)	(0.01)	(0.10)	(0.10)
Market Size*Wins Above Threshold	0.28**	0.30**	0.35**	0.35	0.35
	(0.08)	(0.05)	(0.10)	(0.19)	(0.19)
New Stadium Dummy	36.19**	36.13**	36.12**	36.13**	36.13**
	(6.10)	(6.08)	(5.93)	(5.93)	(5.93)
Stadium Age	5.52**	5.52**	5.57**	5.57**	5.57**
	(0.79)	(0.80)	(0.81)	(0.81)	(0.81)

NOTE: Robust standard errors are given in parentheses below estimates.
For specification 5 market size, wins, and wins above the threshold are centered about the mean.
*Significance at the 5% level.
**Significance at the 1% level.

Copy of Regression Table from "Market Size, Pay, and Performance: A General Model and Application to Major League Baseball" ***Journal of Sports Economics*****, Burger and Walters (2003)**

TABLE 3: Regression Results

	Team Revenue (TR)			
Dependent Variable	*1*	*2*	*3*	*4*
Adjusted R^2	0.709	0.707	0.707	0.705
Number of observations	144	144	144	144
Constant	10.479	9.164	–19.431	–14.823
	(6.308)	(6.447)	(17.449)	(40.251)
Trend	4.933**	4.770**	5.008**	4.999**
	(1.244)	(1.263)	(1.258)	(1.267)
M	–6.088			–1.111
	(3.339)			(7.625)
W			0.392	0.333
			(0.226)	(0.535)
W_2			–0.660	–0.576
			(0.676)	(1.027)
MW	0.132**	0.058**	0.053**	0.067
	(0.046)	(0.017)	(0.018)	(0.106)
MW_2	0.198*	0.301**	0.322**	0.303
	(0.097)	(0.084)	(0.112)	(0.200)
Stadium	40.902**	41.5147**	41.094**	41.084**
	(8.592)	(8.668)	(8.942)	(8.952)
Age	–2.348	–2.324	–2.422	–2.415
	(1.800)	(1.795)	(1.820)	(1.822)

NOTE: Newey-West (1987) standard errors in parentheses.
*Significance at the 5% level. **Significance at the 1% level.

Table 4: Stage 1 Regression Pooling All Teams

	Team Revenue				
Dependent Variable	1	2	3	4	Full
Adjusted R-Square	0.5735	0.5927	0.6281	0.8002	0.8019
Number of Observations	458	458	458	458	458
Constant	19.83	9.89	30.14	-83.70**	-50.16**
	(17.76)	(17.31)	(16.68)	(10.76)	(14.42)
Time Trend	8.22**	8.30**	8.17**	5.58**	5.56**
	(0.40)	(0.38)	(0.36)	(0.28)	(0.28)
Winning Percentage	119.29**	133.41**	84.09*	127.09**	52.4
	(39.49)	(38.21)	(36.83)	(16.81)	(29.16)
Square Root of Win Percent Above Threshold	102.29**	89.47**	77.78**		48.01**
	(27.17)	(26.62)	(25.05)		(18.54)
Inverse of Team Age		-36.35	-40.2	17.36	14.69
		(27.43)	(28.04)	(22.98)	(24.29)
(Inverse of Team Age)^2		68.15*	73.23*	11.98	11.76
		(28.70)	(30.02)	(28.49)	(29.89)
Inverse of Stadium Age		52.67*	43.19	20.39	18.38
		(23.48)	(22.42)	(16.45)	(16.57)
(Inverse of Stadium Age)^2		-17.5	-7.38	-10.83	-8.88
		(29.11)	(29.57)	(17.66)	(17.76)
Lag of Playoffs Dummy Variable			27.05**	17.31**	17.00**
			(4.77)	(3.43)	(3.40)
Price				0.72**	0.71**
				(0.05)	(0.05)

NOTE:
Robust standard errors in parentheses below estimates.
Square Root of Win Percent Above Threshold variable is the square root of:
(team's winning percentage - .506).
If a team has a winning percentage beneath .506 the variable takes a value of 0.
Team Age and Stadium Age take values 0 to 6, after a team or stadium has existed for 6 years its age is considered 0 thereafter.
The inverse of these age variables is 1/(Years of Existence).
*Significant at 5% level.
**Significant at 10% level.

Table 5: Stage 1 Regression Results for Individual Team Effects

Dependent Variable	Team Revenue 1	2	3	Linear Win Effect Only	Full Specification	Lag of Win Percent Included	2 Lags of Win Percent Included
Adjusted R-Square	0.984	0.9883	0.9888	0.9928	0.9929	0.9945	0.9951
Number of Observations	458	458	458	458	458	429	400
Time Trend	8.53**	8.79**	8.73**	6.54**	6.44**	6.15**	6.03**
	(0.31)	(0.27)	(0.27)	(0.25)	(0.26)	(0.27)	(0.29)
Inverse of Team Age		93.45**	88.70**	73.10*	67.26*	97.50*	155.71
		(28.82)	(30.30)	(28.48)	(26.29)	(43.41)	(89.08)
(Inverse of Team Age)^2		-51.26	-46.13	-44.36	-39.42	-143.94	-427.93
		(32.15)	(33.20)	(28.72)	(26.37)	(80.22)	(262.66)
Inverse of Stadium Age		86.45**	83.15**	45.32**	39.75**	44.70**	47.20**
		(16.52)	(15.89)	(13.24)	(13.39)	(12.94)	(13.31)
(Inverse of Stadium Age)^2		-50.22**	-46.54**	-31.17*	-25.57*	-27.64*	-30.36*
		(17.25)	(17.05)	(12.83)	(12.89)	(12.43)	(12.94)
Lag of Playoffs Dummy Variable			9.33**	7.77**	7.31**	1.45	1.32
			(2.77)	(2.06)	(2.07)	(2.15)	(2.12)
Price				0.51**	0.52**	0.50**	0.48**
				(0.04)	(0.04)	(0.04)	(0.04)
Lag of Win Percent						25.22	34.00*
						(16.59)	(15.79)
Second Lag of Win Percent							22.16
							(13.56)

NOTE:
Robust standard errors in parentheses below estimates.
WPT stands for Win Percent Threshold and indicates variable that is the square root of:
(team's winning percentage - .506).
If a team has a winning percentage beneath .506 the variable takes a value of 0.
Team Age and Stadium Age take values 0 to 6, after a team or stadium has existed for 6 years its age is considered 0 thereafter.
The inverse of these age variables is 1/(Years of Existence).
*Significant at 5% level.
**Significant at 10% level.

Table 5 Continued: Stage 1 Regression Results for Individual Team Effects

	Team Revenue						
Dependent Variable	1	2	3	Linear Win Effect Only	Full Specification	Lag of Win Percent Included	2 Lags of Win Percent Included
ARI	110.74*	69.65	83.97*	48.86**	52.52**	40.84	18.81
	(44.46)	(39.75)	(35.88)	(13.33)	(16.64)	(21.27)	(24.95)
ATL	-363.26	-255.1	-204.39	1.65	-264.07*	-309.50**	-256.13**
	(204.42)	(146.92)	(137.37)	(36.33)	(129.59)	(111.99)	(83.60)
BAL	-179.06	-165.54	-152.19	-119.74**	-133.33	-170.8	-166.62
	(119.85)	(130.89)	(124.27)	(32.12)	(89.90)	(92.06)	(85.09)
BOS	8.72	-5.35	-179.4	29.53	-53.91	-3.11	-80.26
	(504.13)	(482.18)	(460.69)	(30.23)	(248.16)	(204.74)	(196.05)
CHC	115.22*	113.17*	104.28	4.57	79.87*	77.62*	74.93
	(55.70)	(50.92)	(53.52)	(15.62)	(31.68)	(38.02)	(45.63)
CIN	128.45*	43.97	63.16	2.53	6.38	-11.5	-33.27
	(58.83)	(40.28)	(34.99)	(25.95)	(29.17)	(32.55)	(46.24)
CLE	76.87	96.74	107.42	-62.55	26.48	-71.96	-102.94
	(133.97)	(149.35)	(141.51)	(55.63)	(113.47)	(64.98)	(57.49)
COL	-106.82	19.96	34.79	36.13	-49.13	-34.19	-40.19
	(100.44)	(62.84)	(71.90)	(29.83)	(48.15)	(40.31)	(38.70)
CWS	193.03	150.61	164.25	-23.46	92.58	105.03	93.14
	(123.24)	(142.55)	(150.10)	(35.40)	(112.58)	(103.93)	(101.77)
DET	50.17	28.16*	29.57*	-60.28**	-51.68**	-56.18**	-62.84**
	(27.36)	(12.34)	(12.22)	(14.20)	(16.13)	(18.13)	(22.26)
FLA	102.82**	84.35**	66.15**	-52.65	1.6	-1.59	-16.43
	(33.92)	(20.57)	(20.34)	(42.62)	(18.30)	(19.51)	(27.26)
HOU	154.4	-13.28	9.3	-18.12	-98.39	-103.17	-129.93
	(135.63)	(110.67)	(98.44)	(27.07)	(68.84)	(79.97)	(95.02)
KC	47.26**	40.25*	41.91*	8.73	28.4	6.73	-10.06
	(17.11)	(17.65)	(17.36)	(23.82)	(24.45)	(21.23)	(21.93)
LAA	-23.82	-13.95	13.37	-103.09**	-162.4	-179.68	-103.74
	(106.97)	(102.42)	(95.61)	(33.67)	(124.52)	(137.70)	(116.87)
LAD	87.7	95.77	99.94*	52.22*	150.68**	126.22**	94.81
	(60.12)	(55.06)	(46.24)	(23.17)	(42.18)	(38.04)	(48.74)
MIL	80.93**	-31.3	-22.35	-87.04**	-62.14**	-71.17**	-84.93**
	(24.62)	(20.09)	(19.86)	(19.66)	(17.86)	(20.22)	(20.56)
MIN	-11.4	-13.19	10.52	-49.71**	-56.91*	-79.12**	-90.62*
	(36.96)	(34.71)	(37.48)	(9.90)	(24.00)	(30.11)	(43.15)
NYM	82.79	69.78	91.08	-8.96	-10.08	-44.79	-18.62
	(131.90)	(152.85)	(136.21)	(25.62)	(68.54)	(76.98)	(72.89)
NYY	869.53	720.33	691.9	46.23	634.45	405.15	442.22
	(557.84)	(480.30)	(469.08)	(111.20)	(350.35)	(353.32)	(406.19)
OAK	126.68**	135.00**	145.81**	-41.4	92.96**	62.60**	54.33*
	(44.87)	(48.91)	(46.14)	(25.72)	(25.27)	(22.91)	(24.85)
PHI	59.11**	63.09**	60.37**	-79.24**	-26.65	-18.75	-19.41
	(17.24)	(16.60)	(15.43)	(16.08)	(30.65)	(30.02)	(31.17)
PIT	18.1	-0.3	0.89	8.13	8.1	26.84	21.46
	(53.15)	(23.53)	(22.88)	(30.60)	(32.31)	(31.25)	(32.67)

Table 5 Continued: Stage 1 Regression Results for Individual Team Effects

Dependent Variable	Team Revenue 1	2	3	Linear Win Effect Only	Full Specification	Lag of Win Percent Included	2 Lags of Win Percent Included
SD	30.88	40.67	45.2	-64.81**	-41.32	-52.08**	-63.50**
	(52.37)	(25.98)	(28.64)	(18.92)	(22.56)	(18.66)	(22.70)
SEA	78.24	99.95	116.39*	8.31	50.32	43.18	26
	(68.26)	(54.12)	(57.30)	(20.17)	(42.86)	(44.13)	(46.13)
SF	-75.96	-69.19	-70.46	-21.9	-142.15	-241.86	65.05
	(165.62)	(152.80)	(155.69)	(37.35)	(149.14)	(134.63)	(107.09)
STL	-214.25**	-33.28	-14.44	-0.38	-13.14	-1.93	-30.25
	(67.56)	(158.53)	(176.27)	(28.47)	(134.36)	(191.85)	(204.04)
TB	86.99	47.72	64.57	-55.66*	-0.54	-32.3	-57.28
	(52.27)	(37.54)	(45.12)	(24.07)	(60.50)	(64.00)	(72.43)
TEX	290.92**	346.96**	329.25**	16.67	27.57	5.69	-0.57
	(106.04)	(107.17)	(91.92)	(16.16)	(47.37)	(34.42)	(34.90)
TOR	132.96	94.49	95.02	42.72	12.61	-63.84	-115.89**
	(78.78)	(62.26)	(62.13)	(45.84)	(58.00)	(87.05)	(25.03)

Table 5 Continued: Stage 1 Regression Results for Individual Team Effects

Dependent Variable	Team Revenue 1	2	3	Linear Win Effect Only	Full Specification	Lag of Win Percent Included	2 Lags of Win Percent Included
ARI Win Percent	-79.8	-18.47	-55.3	-79.76**	-90.46*	-71.7	-44.02
	(83.49)	(87.18)	(81.04)	(28.60)	(36.80)	(40.73)	(47.66)
ATL Win Percent	968.57*	734.48*	620.13*	56.09	637.09*	725.78**	585.06**
	(427.72)	(306.32)	(290.07)	(66.22)	(281.83)	(242.28)	(181.36)
BAL Win Percent	619.61*	580.93	550.69*	321.07**	349.2	429.61*	398.22*
	(268.77)	(296.22)	(279.95)	(66.69)	(206.33)	(213.01)	(197.01)
BOS Win Percent	225.71	253.22	610.28	-24.69	146.25	27.39	157.17
	(1045.32)	(999.75)	(954.63)	(57.69)	(517.00)	(428.76)	(408.10)
CHC Win Percent	-16.27	-15.4	3.13	22.82	-155.30*	-160.68	-180.04
	(129.55)	(118.29)	(122.60)	(30.84)	(73.43)	(91.44)	(112.31)
CIN Win Percent	-148.26	5.7	-35.9	-17.54	-27.7	-2.54	21.24
	(124.60)	(84.87)	(72.18)	(54.52)	(62.05)	(65.57)	(97.30)
CLE Win Percent	17.76	-30.67	-60.49	164.59	-39.67	160.39	200.44
	(279.63)	(311.42)	(296.60)	(110.69)	(238.56)	(138.14)	(123.41)
COL Win Percent	446.35*	142.95	108.34	-37.69	149.53	104.94	95.32
	(226.80)	(138.16)	(159.10)	(60.97)	(106.78)	(86.69)	(82.99)
CWS Win Percent	-238.26	-158	-189.31	40.92	-206.82	-242.59	-241.3
	(255.06)	(295.08)	(311.49)	(65.22)	(235.54)	(214.09)	(208.87)
DET Win Percent	44.65	68.66*	66.45*	121.21**	95.73*	95.37*	90.57
	(71.39)	(31.70)	(31.59)	(30.49)	(37.92)	(41.21)	(49.68)
FLA Win Percent	-130.29	-104.69*	-66.5	75.48	-52.53	-59.5	-54.01
	(80.17)	(49.61)	(49.61)	(90.74)	(42.24)	(44.18)	(62.20)
HOU Win Percent	-143.18	193.25	141.26	40.88	210.49	211.58	246.02
	(287.11)	(233.78)	(203.97)	(46.25)	(139.04)	(162.22)	(197.19)
KC Win Percent	-9.88	2.37	-0.53	-48.96	-100.88	-57.71	-38.64
	(40.06)	(42.07)	(41.40)	(58.36)	(60.79)	(47.87)	(47.29)
LAA Win Percent	212.92	188.95	126.19	242.11**	367.78	394.9	214.71
	(223.72)	(213.41)	(197.28)	(60.48)	(272.43)	(299.82)	(255.87)
LAD Win Percent	76.34	54.72	37.16	-18.16	-230.64*	-186.46*	-144.44
	(127.87)	(118.24)	(101.52)	(41.55)	(93.87)	(82.10)	(103.22)
MIL Win Percent	-59.45	163.31**	143.45**	175.74**	116.90**	125.52**	134.48**
	(51.37)	(41.66)	(41.19)	(42.02)	(39.03)	(44.82)	(46.41)
MIN Win Percent	118.6	121.89	64.24	59.60**	73.11	111.74	114.85
	(84.48)	(77.19)	(84.98)	(18.62)	(57.46)	(70.84)	(104.75)
NYM Win Percent	88.04	95.46	47.62	86.81	85.14	155.09	81.51
	(291.61)	(347.12)	(308.85)	(47.77)	(155.47)	(176.32)	(169.26)
NYY Win Percent	-1660.13	-1296.82	-1244.39	110.59	-1314.53	-691.68	-786.11
	(1407.61)	(1234.37)	(1196.74)	(177.63)	(898.38)	(889.25)	(1029.09)
OAK Win Percent	-163.66	-185.83	-209.61*	54.26	-243.81**	-189.42**	-196.90**
	(102.96)	(112.12)	(105.35)	(45.38)	(57.26)	(47.92)	(53.53)
PHI Win Percent	2.86	-12.93	-4.41	158.20**	34.85	-2.4	-21.29
	(42.05)	(39.17)	(36.60)	(31.45)	(69.15)	(69.45)	(70.01)
PIT Win Percent	69.72	88.35	86.99	-57.17	-59.31	-112.97	-122.3
	(117.93)	(53.13)	(51.39)	(71.91)	(75.54)	(72.54)	(75.30)

Table 5 Continued: Stage 1 Regression Results for Individual Team Effects

Dependent Variable	Team Revenue 1	2	3	Linear Win Effect Only	Full Specification	Lag of Win Percent Included	2 Lags of Win Percent Included
SD Win Percent	67.89	26.11	12.81	116.44**	60.44	75.67*	76.59
	(122.60)	(61.01)	(66.36)	(35.62)	(49.05)	(36.48)	(44.49)
SEA Win Percent	29.97	-36.61	-75.78	16.42	-87.74	-79.52	-60.15
	(153.21)	(121.92)	(130.69)	(39.04)	(98.16)	(96.53)	(104.68)
SF Win Percent	354.2	339.12	342.77	55.61	320.23	541.32	-150.81
	(371.87)	(342.96)	(349.26)	(67.96)	(329.03)	(293.60)	(240.77)
STL Win Percent	666.43**	243.9	193.74	12.13	37.41	5.82	39.73
	(146.54)	(352.17)	(395.16)	(50.79)	(297.93)	(411.40)	(437.67)
TB Win Percent	-74.38	-18.14	-59.28	85.88	-54.98	14.88	60.28
	(123.79)	(90.61)	(110.45)	(45.66)	(142.13)	(150.33)	(171.65)
TEX Win Percent	-429.74	-562.02*	-526.26**	11.18	-15.25	24.13	13.47
	(225.78)	(226.69)	(194.62)	(29.95)	(99.71)	(70.25)	(71.71)
TOR Win Percent	-149.05	-72.26	-72.46	-100.12	-35.19	112	189.37**
	(164.77)	(130.34)	(130.08)	(92.09)	(120.13)	(174.75)	(45.12)

Table 5 Continued: Stage 1 Regression Results for Individual Team Effects

Dependent Variable	Team Revenue 1	2	3	Linear Win Effect Only	Full Specification	Lag of Win Percent Included	2 Lags of Win Percent Included
ARI WPT	115.49**	-13.38	5.16		14.33	3.07	-21.74
	(35.69)	(50.17)	(52.54)		(26.64)	(25.26)	(33.20)
ATL WPT	-317.36	-251.77	-213.91		-283.58*	-313.98**	-248.71**
	(182.63)	(130.98)	(127.96)		(131.70)	(118.89)	(88.12)
BAL WPT	-97.41	-117.21	-120.9		-20.88	-77.77	-58.69
	(139.79)	(158.36)	(147.40)		(112.83)	(116.27)	(105.11)
BOS WPT	12.61	-2.3	-143.97		-68.41	-3.55	-35.39
	(372.73)	(355.91)	(338.94)		(194.54)	(155.28)	(145.20)
CHC WPT	54.55	51.25	27.3		115.09*	133.99*	156.26*
	(81.07)	(73.81)	(72.87)		(48.74)	(56.96)	(70.80)
CIN WPT	16.25	-4.3	15.37		4.15	-20.32	-31.79
	(57.77)	(44.62)	(38.88)		(40.62)	(55.07)	(68.35)
CLE WPT	120.94	104.83	115.46		115.83	71.49	78.28
	(117.31)	(122.61)	(119.74)		(96.41)	(77.61)	(76.15)
COL WPT	-271.84*	-185.80*	-168.02		-111.18	-61.39	-54.12
	(119.33)	(82.86)	(91.45)		(58.19)	(39.59)	(40.09)
CWS WPT	111	84.11	96.85		101.1	116.14	118.67
	(93.15)	(106.65)	(113.06)		(92.62)	(80.84)	(80.25)
DET WPT	32.66	57.50*	47.05		34.93	45.16	44.55
	(60.97)	(27.04)	(32.26)		(36.38)	(35.46)	(38.16)
FLA WPT	131.39	120.54	106.22		118.18	120.83	125.06
	(102.28)	(74.86)	(75.76)		(72.67)	(65.11)	(66.77)
HOU WPT	71.5	-75.13	-61.41		-75.79	-77.04	-97.92
	(124.11)	(100.93)	(80.39)		(48.80)	(61.85)	(80.87)
KC WPT	73.22	56.53	60.49		241.16**	183.70**	162.61*
	(63.30)	(67.53)	(66.32)		(92.94)	(66.51)	(65.09)
LAA WPT	-20.36	-14	5.76		-56.55	-55.35	5.44
	(90.28)	(84.67)	(73.81)		(132.32)	(144.35)	(129.47)
LAD WPT	-46.29	-36.04	-21.39		86.93	69.15	54.08
	(54.41)	(52.92)	(53.27)		(45.33)	(40.53)	(48.92)
MIL WPT	83.89**	17.52	31.69		79.50**	74.34*	65.59
	(26.73)	(21.59)	(23.48)		(29.82)	(35.60)	(35.45)
MIN WPT	-18	-52.17	-29.47		-6.44	-21.91	-29.94
	(57.59)	(37.93)	(45.72)		(36.02)	(44.70)	(63.99)
NYM WPT	9.36	38.36	54.09		0.64	-39.19	-19.14
	(151.56)	(191.96)	(170.71)		(87.98)	(101.02)	(99.26)
NYY WPT	1171.19	921.05	885.23		884.41	409.28	440.66
	(974.40)	(872.80)	(838.79)		(633.63)	(614.23)	(714.37)
OAK WPT	100.97	111.06	104.41		149.53**	125.43**	123.11**
	(55.90)	(61.10)	(57.27)		(32.67)	(26.56)	(30.12)
PHI WPT	191.78**	135.29**	111.43**		75.23	119.44**	119.50**
	(26.98)	(30.66)	(27.18)		(42.93)	(44.15)	(44.33)
PIT WPT	0	0	0		0	0	0
	-	-	-		-	-	-
SD WPT	4.73	8.5	12.14		36.85	24.73	25.96
	(82.76)	(54.46)	(56.21)		(33.13)	(23.22)	(27.97)

Table 5 Continued: Stage 1 Regression Results for Individual Team Effects

	Team Revenue						
Dependent Variable	1	2	3	Linear Win Effect Only	Full Specification	Lag of Win Percent Included	2 Lags of Win Percent Included
SEA WPT	74.29	86.93	100.15		72.31	75.08	60.51
	(94.03)	(75.97)	(82.20)		(65.50)	(62.81)	(71.78)
SF WPT	-87.39	-143.8	-159.58		-130.64	-257.78	50.38
	(190.61)	(174.24)	(178.06)		(158.30)	(132.45)	(118.79)
STL WPT	-304.61**	-85.56	-61.3		-12.64	-0.45	-17.62
	(68.60)	(181.82)	(205.91)		(153.15)	(187.89)	(198.29)
TB WPT	0.77	11.46	35.56		108.31	61.76	18.39
	(78.63)	(62.76)	(74.24)		(86.46)	(88.92)	(105.46)
TEX WPT	210.89*	234.58*	219.33**		15.05	-4.75	-1.42
	(103.03)	(92.33)	(79.44)		(45.49)	(31.64)	(32.86)
TOR WPT	51.82	21.91	22.69		-48.51	-95.25	-91.35
	(75.78)	(70.12)	(69.70)		(69.97)	(74.01)	(55.23)

Table 6: Summary Statistics Across Teams for Stage 2 Regression Variables

Variable	Mean	Number of Observations	SE	Minimum	Maximum
Linear Combination of Estimates (Specification 5)	48.56	29	156.93	-430.12	353.52
Team Dummy Variable Coefficient Estimate (Specification 5)	3.40	29	148.55	-264.07	634.45
Win Percentage Coefficient Estimate (Specification 4)	67.17	29	109.72	-166.58	389.05
Population (Millions)	6.75	29	5.56	1.71	21.52
Per Capita Income (Real 2010 Dollars)	42734.93	29	5470.61	30678.73	55115.29
Number of Other Major Sports Franchises	3.11	29	1.77	1.00	8.00
Distance to Nearest MLB Team (Miles)	170.12	29	165.09	7.00	679.00
Division Category	3.55	29	1.78	1.00	6.00

SOURCE:
Population and per capita income data for teams in the United States from U.S. Census and the Bureau of Economic Analysis, respectively.
Population data for TOR from Statistics Canada. Per capita income data for TOR for the years 2001-2010 from the Montreal Chamber of Commerce and is linearly extrapolated backwards for the years 1995-2000.
Per capita income for TOR is converted to U.S. dollars based on the OANDA yearly average exchange rate.
All dollar figures are in real 2010 U.S. dollars using consumer price index (CPI) price deflator.

Other sports franchise data from hockey-reference.com, basketball-reference.com, and pro-football-reference.com.
Only teams within the NHL, NBA, or NFL in the same metropolitan area are considered additional sports franchises.
Distance to nearest MLB team data from sportmapworld.com.
Division indicator is as follows:

AL East=1	AL Central=2	AL West=3
NL East=4	NL Central=5	NL West=6

Table 7: Time Average of Stage 2 Regression Variables by Team

Team	Linear Combination of Estimates (Specification 5)	Standard Error of Linear Combination of Estimates	Team Dummy Variable Coefficient Estimate (Specification 5)	Standard Error of Team Dummy Variable Estimate	Win Percentage Coefficient Estimate (Specification 4)
ARI	-76.13	27.34	52.52	16.64	-79.76
ATL	353.52	156.33	-264.07	129.59	56.09
BAL	328.32	98.42	-133.33	89.90	321.07
BOS	77.84	324.42	-53.91	248.16	-24.69
CHC	-40.21	31.54	79.87	31.68	22.82
CIN	-23.55	46.20	6.38	29.17	-17.54
CLE	76.16	158.98	26.48	113.47	164.59
COL	38.35	67.20	-49.13	48.15	-37.69
CWS	-105.72	146.47	92.58	112.58	40.92
DET	130.65	26.82	-51.68	16.13	121.21
FLA	65.65	61.40	1.60	18.30	75.48
HOU	134.70	93.91	-98.39	68.84	40.88
KC	140.28	36.68	28.40	24.45	-48.96
LAA	311.23	146.42	-162.40	124.52	242.11
LAD	-143.71	52.87	150.68	42.18	-18.16
MIL	196.40	35.77	-62.14	17.86	175.74
MIN	66.67	26.10	-56.91	24.00	59.60
NYM	85.78	76.05	-10.08	68.54	86.81
NYY	-430.12	276.75	634.45	350.35	110.59
OAK	-94.28	31.66	92.96	25.27	54.26
PHI	110.08	35.59	-26.65	30.65	158.20
PIT	-59.31	75.54	8.10	32.31	-57.17
SD	97.29	29.14	-41.32	22.56	116.44
SEA	-15.43	46.20	50.32	42.86	16.42
SF	189.59	176.68	-142.15	149.14	55.61
STL	24.78	148.41	-13.14	134.36	12.13
TB	53.33	62.44	-0.54	60.50	85.88
TEX	-0.20	58.25	27.57	47.37	11.18
TOR	-83.70	94.27	12.61	58.00	-100.12
Total	48.56	91.31	3.40	75.09	56.69

SOURCE:

Population and per capita income data for teams in the United States from U.S. Census and the Bureau of Economic Analysis, respectively.

Population data for TOR from Statistics Canada. Per capita income data for TOR for the years 2001-2010 from the Montreal Chamber of Commerce and is linearly extrapolated backwards for the years 1995-2000.

Per capita income for TOR is converted to U.S. dollars based on the OANDA yearly average exchange rate.

All dollar figures are in real 2010 U.S. dollars using consumer price index (CPI) price deflator.

Table 7 Continued: Time Average of Stage 2 Regression Variables by Team

Team	Standard Error of Win Percentage Estimate	Population (Millions)	Per Capita Income (Real 2010 Dollars)	Number of Other Major Sports Franchises	Distance to Nearest MLB Team (Miles)	Division Category
ARI	16.64	3.67	37253.41	3.00	299.00	6
ATL	129.59	4.90	39212.63	2.75	371.00	4
BAL	89.90	7.84	49529.25	4.31	68.75	1
BOS	248.16	7.32	46787.87	3.00	180.00	1
CHC	31.68	9.42	43723.90	4.00	8.00	5
CIN	29.17	2.10	38517.93	1.00	223.00	5
CLE	113.47	2.92	38714.44	1.81	91.00	2
COL	48.15	2.75	45429.83	3.00	564.00	6
CWS	112.58	9.42	43723.90	4.00	8.00	2
DET	16.13	5.36	39570.68	3.00	91.00	2
FLA	18.30	5.15	41255.98	3.00	269.63	4
HOU	68.84	5.16	42342.02	1.69	231.00	5
KC	24.45	1.97	39472.66	1.00	233.00	2
LAA	124.52	16.80	38952.48	4.00	28.00	3
LAD	42.18	16.80	38952.48	4.00	28.00	6
MIL	17.86	1.71	41057.38	1.00	76.00	5
MIN	24.00	3.36	44761.65	2.69	297.00	2
NYM	68.54	21.52	50379.64	8.00	7.00	4
NYY	350.35	21.52	50379.64	8.00	7.00	1
OAK	25.27	7.10	55115.29	5.00	10.00	3
PHI	30.65	6.31	42978.82	3.00	89.00	4
PIT	32.31	2.50	39045.16	2.00	114.00	5
SD	22.56	2.87	42921.64	1.00	86.00	6
SEA	42.86	3.82	45566.03	1.88	679.00	3
SF	149.14	7.10	55115.29	5.00	10.00	6
STL	134.38	2.80	39706.20	2.00	233.00	5
TB	60.50	2.58	37519.89	2.00	195.00	1
TEX	47.37	5.84	40648.12	3.00	231.00	3
TOR	58.00	5.02	30678.73	2.00	206.00	1
Total	75.09	6.75	42734.93	3.11	170.12	3.56

Other sports franchise data from hockey-reference.com, basketball-reference.com, and pro-football-reference.com.

Only teams within the NHL, NBA, or NFL in the same metropolitan area are considered additional sports franchises.

Distance to nearest MLB team data from sportmapworld.com.

Division indicator is as follows:

AL East=1	AL Central=2	AL West=3
NL East=4	NL Central=5	NL West=6

Table 8: Stage 2 Regression Results for Team Intercepts from Stage 1 Regression

	Team Intercepts from Stage 1					
Dependent Variable	1	2	3	4	5	6
R-Square*	0.4368	0.4099	0.4264	0.4472	0.2269	0.4522
Number of Observations	29	29	29	29	29	29
Constant	103.292	-89.957**	59.411	109.085	-47.984	81.802
	(1.182)	(-3.901)	(0.675)	(1.367)	(-0.726)	(0.917)
Population		5.465	8.087**	3.244	4.701	4.158
		(1.607)	(3.410)	(1.016)	(1.380)	(1.201)
Per Capita Income	-0.005*		-0.003	-0.005*	-0.001	-0.004*
	(-2.261)		(-1.606)	(-2.327)	(-0.342)	(-1.992)
Pro Sports Franchises	22.497**	8.973		15.161	13.144	14.4
	(3.551)	(1.015)		(1.650)	(1.491)	(1.556)
Distance to Closest Team	0.012	0.074	0.049		0.078	0.039
	(0.237)	(1.383)	(0.877)		(1.579)	(0.683)
AL East Dummy	2.608	16.315	2.715	0.684		1.589
	(0.065)	(0.417)	(0.068)	(0.017)		(0.040)
AL West Dummy	109.867**	71.779**	110.785**	113.458**		107.500**
	(3.605)	(2.903)	(3.636)	(3.876)		(3.520)
NL East Dummy	11.212	1.303	8.048	10.546		6.947
	(0.557)	(0.064)	(0.394)	(0.534)		(0.340)
NL Central Dummy	33.284	31.514	24.624	29.389		32.21
	(1.811)	(1.713)	(1.388)	(1.639)		(1.751)
NL West Dummy	51.053**	50.933**	48.190**	50.817**		48.768**
	(2.962)	(2.943)	(2.779)	(2.975)		(2.812)

NOTE:
t-statistic in parentheses beneath estimate.
Dependent variable is the estimate of the team dummy variable coefficients from the full specification of the stage 1 regression.
Explanatory variables are the time average team means for population, per capita income, number of other pro sports franchises, the distance in miles to the closest MLB team, and dummy variables for each MLB division omitting the AL Central.
The model is estimated using variance weighted least squares (WLS) based on the standard error of the dependent variable from the stage 1 regression.
*R-Square values are obtained from estimating the model using ordinary least squares (OLS) weighting each observation by the inverse of the estimate's variance.

Table 9: Stage 2 Regression Results for Linear Combination of Team Estimates of Win Effects

	Linear Combination of Stage 1 Team Estimates					
Dependent Variable	1	2	3	4	5	6
R-Square*	0.6168	0.5749	0.4877	0.5903	0.3075	0.6179
Number of Observations	29	29	29	29	29	29
Constant	9.61	240.895**	127.356	-91.762	83.742	-7.541
	(0.083)	(7.518)	(1.105)	(-0.852)	(0.916)	(-0.063)
Population		-0.11	-11.858**	4.836	-0.286	2.543
		(-0.025)	(-4.077)	(1.132)	(-0.066)	(0.565)
Per Capita Income	0.006*		0.001	0.008**	0.002	0.006*
	(2.080)		(0.450)	(2.732)	(1.038)	(2.155)
Pro Sports Franchises	-48.621**	-42.584**		-55.445**	-42.450**	-54.196**
	(-5.820)	(-3.623)		(-4.296)	(-3.382)	(-4.191)
Distance to Closest Team	-0.127	-0.158*	-0.132		-0.151*	-0.114
	(-1.894)	(-2.333)	(-1.874)		(-2.354)	(-1.620)
AL East Dummy	-16.203	-33.68	-26.838	-10.227		-15.873
	(-0.333)	(-0.702)	(-0.552)	(-0.211)		(-0.326)
AL West Dummy	-129.912**	-90.119**	-130.292**	-149.180**		-132.268**
	(-3.577)	(-2.918)	(-3.564)	(-4.258)		(-3.618)
NL East Dummy	39.926	41.405	39.296	34.199		36.095
	(1.184)	(1.207)	(1.143)	(0.995)		(1.050)
NL Central Dummy	-84.661**	-87.601**	-63.497*	-76.803**		-87.054**
	(-3.146)	(-3.215)	(-2.382)	(-2.898)		(-3.195)
NL West Dummy	-107.739**	-114.561**	-98.199**	-110.943**		-110.085**
	(-4.389)	(-4.618)	(-3.970)	(-4.457)		(-4.422)

NOTE:
t-statistic in parentheses beneath estimate.
Dependent variable is the linear combination of the team estimate for the coefficient of the win percentage variable and the square root of win percentage above threshold variable from the full specification of the stage 1 regression.
Explanatory variables are the time average team means for population, per capita income, number of other pro sports franchises, the distance in miles to the closest MLB team, and dummy variables for each MLB division omitting the AL Central.
The model is estimated using variance weighted least squares (WLS) based on the standard error of the dependent variable from the stage 1 regression.
*R-Square values are obtained from estimating the model using ordinary least squares (OLS) weighting each observation by the inverse of the estimate's variance.

Table 10: Stage 2 Regression Results for Linear Team Win Percent Effects

	Stage 1 Linear Win Percent Team Estimates					
Dependent Variable	1	2	3	4	5	6
R-Square*	0.5509	0.4719	0.4845	0.417	0.38	0.5525
Number of Observations	29	29	29	29	29	29
Constant	-152.637	194.650**	-108.288	-378.208**	-142.611*	-163.133
	(-1.746)	(8.440)	(-1.230)	(-4.740)	(-2.159)	(-1.828)
Population		-0.691	-7.591**	9.236**	3.008	2.031
		(-0.203)	(-3.201)	(2.892)	(0.883)	(0.586)
Per Capita Income	0.009**		0.006**	0.013**	0.007**	0.009**
	(4.114)		(3.169)	(6.166)	(4.536)	(4.150)
Pro Sports Franchises	-31.315**	-23.966**		-41.274**	-30.141**	-35.270**
	(-4.943)	(-2.710)		(-4.493)	(-3.418)	(-3.812)
Distance to Closest Team	-0.317**	-0.379**	-0.331**		-0.287**	-0.304**
	(-6.080)	(-7.053)	(-5.883)		(-5.786)	(-5.380)
AL East Dummy	13.11	-18.064	9.853	19.744		12.612
	(0.329)	(-0.462)	(0.247)	(0.496)		(0.317)
AL West Dummy	-53.408	19.845	-62.611*	-101.528**		-54.565
	(-1.752)	(0.803)	(-2.055)	(-3.469)		(-1.787)
NL East Dummy	54.736**	64.411**	49.958*	24.281		52.653*
	(2.719)	(3.182)	(2.445)	(1.229)		(2.576)
NL Central Dummy	-37.086*	-36.161*	-19.029	-15.368		-37.611*
	(-2.018)	(-1.966)	(-1.073)	(-0.857)		(-2.044)
NL West Dummy	-52.119**	-57.745**	-51.819**	-69.386**		-53.235**
	(-3.024)	(-3.336)	(-2.989)	(-4.062)		(-3.070)

NOTE:
t-statistic in parentheses beneath estimate.
Dependent variable is the estimated coefficient on the team win percentage variable in specification 4 of the stage 1 regression.
Explanatory variables are the time average team means for population, per capita income, number of other pro sports franchises, the distance in miles to the closest MLB team, and dummy variables for each MLB division omitting the AL Central.
The model is estimated using variance weighted least squares (WLS) based on the standard error of the dependent variable from the stage 1 regression.
*R-Square values are obtained from estimating the model using ordinary least squares (OLS) weighting each observation by the inverse of the estimate's variance.

Table 11: Stage 1 Regression Results for Teams Categorized Based on Data

	Team Revenue			
Dependent Variable	1	2	3	4
Adjusted R-Square	0.928	0.9381	0.9429	0.9767
Number of Observations	458	458	458	458
Time Trend	11.53**	11.01**	10.69**	5.46**
	(8.110)	(8.064)	(8.196)	(5.606)
Inverse of Team Age		65.75	55.43	8.57
		(1.322)	(1.122)	(0.435)
(Inverse of Team Age)^2		-26.77	-12.04	23.86
		(-0.541)	(-0.244)	(1.141)
Inverse of Stadium Age		182.73**	162.95**	55.26**
		(6.502)	(5.883)	(2.886)
(Inverse of Stadium Age)^2		-137.92**	-117.40**	-36.07
		(-4.641)	(-3.856)	(-1.814)
Lag of Playoffs Dummy Variable			28.89**	17.56**
			(5.414)	(5.089)
Price				0.55**
				(22.971)

NOTE:
t-statistic in parentheses beneath estimate.
Teams are categorized based on Stage 1 regression results with individual team effects.
The categories are as follows:
Category 1: high team intercept, low overall team returns to winning
Category 2: low team intercept, high overal team returns to winning
Category 3: slightly above average team intercept, slightly below average team returns to winning
Category 4: slightly below average team intercept, slightly above average team returns to winning
Category 5: average team intercept, average team returns to winning
These categories are then interacted with time period dummy variables indicating the time period of the three different Collective Bargaining Agreements (CBA).
The coefficients listed are for each category's intercept, win percentage effect, and square root of win percentage above the .506 threshold effect (listed as WPT for win percentage threshold) within the stated time period.
*Signficant at 5% level.
**Significant at 1% level.

Table 11 Continued: Stage 1 Regression Results for Teams Categorized Based on Data

Dependent Variable	Team Revenue 1	2	3	4
Category Estimates within First Time Period:				
Category 1 Intercept	-32.73	-29.72	-49.38	-34.74
	(-0.350)	(-0.315)	(-0.537)	(-0.506)
Category 2 Intercept	-17.89	-99.4	-78.48	-84.19
	(-0.167)	(-0.780)	(-0.673)	(-0.895)
Category 3 Intercept	101.56	-26.32	-2.78	-37.21
	(0.987)	(-0.408)	(-0.052)	(-0.929)
Category 4 Intercept	-203.43*	-104.93	-97.38	-179.87**
	(-2.421)	(-1.945)	(-1.749)	(-3.486)
Category 5 Intercept	-32	-81.55	-88.19	-105.10**
	(-0.368)	(-0.917)	(-1.478)	(-3.015)
Category 1 Win Percent	237.44	234.99	275.16	84.28
	(1.169)	(1.146)	(1.358)	(0.551)
Category 2 Win Percent	188.35	354.94	307.16	188.31
	(0.778)	(1.249)	(1.192)	(0.901)
Category 3 Win Percent	-87.92	173.9	114.64	72.05
	(-0.396)	(1.275)	(1.027)	(0.844)
Category 4 Win Percent	584.25**	351.49**	331.83**	390.07**
	(2.998)	(2.931)	(2.637)	(3.241)
Category 5 Win Percent	208.99	312.45	318.87*	228.79**
	(0.999)	(1.468)	(2.202)	(2.708)
Category 1 WPT	58.79	61.35	-12.8	87.59
	(0.472)	(0.489)	(-0.103)	(0.945)
Category 2 WPT	59.36	-94.42	-129.27	-43.53
	(0.434)	(-0.619)	(-0.952)	(-0.391)
Category 3 WPT	145.67	-67.8	-41.56	-0.2
	(1.204)	(-0.833)	(-0.600)	(-0.004)
Category 4 WPT	-284.70*	-155.91*	-130.65	-159.20*
	(-2.465)	(-2.064)	(-1.666)	(-2.113)
Category 5 WPT	128.67	66.7	30.26	36.2
	(0.917)	(0.475)	(0.298)	(0.588)

Table 11 Continued: Stage 1 Regression Results for Teams Categorized Based on Data

	Team Revenue			
Dependent Variable	1	2	3	4
Category Estimates within Second Time Period:				
Category 1 Intercept	167.5	176.33	167.4	79.62
	(1.373)	(1.443)	(1.351)	(0.878)
Category 2 Intercept	14.18	-88.08	-48.25	-80.98**
	(0.662)	(-1.770)	(-1.176)	(-3.394)
Category 3 Intercept	100.60*	75.78	82.96	35.94
	(2.110)	(1.308)	(1.486)	(0.875)
Category 4 Intercept	12.97	126.46*	131.67*	34.53
	(0.250)	(2.028)	(2.093)	(0.551)
Category 5 Intercept	-46.35	-45.72	-23.05	-32.52
	(-0.905)	(-0.892)	(-0.444)	(-0.907)
Category 1 Win Percent	-230.31	-239.12	-216.49	-165.02
	(-0.880)	(-0.909)	(-0.810)	(-0.849)
Category 2 Win Percent	86.56	289.39*	196.87*	192.82**
	(1.612)	(2.540)	(2.081)	(3.459)
Category 3 Win Percent	-124.88	-96.91	-103.35	-73
	(-1.263)	(-0.778)	(-0.862)	(-0.816)
Category 4 Win Percent	82.03	-183.58	-190.86	-69.77
	(0.681)	(-1.293)	(-1.324)	(-0.505)
Category 5 Win Percent	206.09	209.95	156.3	77.17
	(1.588)	(1.621)	(1.188)	(0.873)
Category 1 WPT	280.65*	284.28*	216.56	187.22
	(2.143)	(2.142)	(1.610)	(1.884)
Category 2 WPT	57.98	-43.96	-41.37	-48.61
	(1.451)	(-0.618)	(-0.700)	(-1.362)
Category 3 WPT	191.69**	146.48	95.52	77.16
	(2.838)	(1.766)	(1.094)	(1.365)
Category 4 WPT	-79.41	12.45	-21.59	-31.04
	(-1.015)	(0.157)	(-0.250)	(-0.441)
Category 5 WPT	-34.8	-34.62	-35.01	-8.27
	(-0.346)	(-0.344)	(-0.340)	(-0.121)

Table 11 Continued: Stage 1 Regression Results for Teams Categorized Based on Data

	Team Revenue			
Dependent Variable	1	2	3	4
Category Estimates within Third Time Period:				
Category 1 Intercept	-266.45	-169.9	-41.25	-183.47
	(-0.646)	(-0.460)	(-0.115)	(-0.711)
Category 2 Intercept	14.11	18.79	93.62	-39.35
	(0.188)	(0.255)	(1.040)	(-0.712)
Category 3 Intercept	-45.05	-25.5	6.54	31.22
	(-0.813)	(-0.455)	(0.115)	(0.914)
Category 4 Intercept	92.80*	-66.78	69.41	-27.66
	(2.366)	(-1.169)	(1.687)	(-0.810)
Category 5 Intercept	-64.04	37.83	110.33	72.3
	(-0.573)	(0.484)	(1.734)	(1.503)
Category 1 Win Percent	640.62	457	169.51	371.24
	(0.741)	(0.588)	(0.226)	(0.689)
Category 2 Win Percent	37.73	43.06	-118.04	115.38
	(0.232)	(0.269)	(-0.595)	(0.924)
Category 3 Win Percent	140.84	103.39	35.82	-50.61
	(1.305)	(0.926)	(0.307)	(-0.727)
Category 4 Win Percent	-179.11*	164.65	-153.54	33.35
	(-2.018)	(1.252)	(-1.689)	(0.432)
Category 5 Win Percent	183.77	-57.89	-222.92	-184.71
	(0.731)	(-0.332)	(-1.613)	(-1.756)
Category 1 WPT	157.39	185.79	277.14	70.98
	(0.461)	(0.580)	(0.981)	(0.315)
Category 2 WPT	57.01	54.13	102.93	22.49
	(0.706)	(0.678)	(1.050)	(0.304)
Category 3 WPT	-20.71	21.6	69.56	41.51
	(-0.305)	(0.303)	(0.930)	(0.755)
Category 4 WPT	204.65**	21.31	163.44*	62.09
	(2.756)	(0.204)	(2.374)	(1.145)
Category 5 WPT	-18.67	120.98	206.68*	171.54*
	(-0.117)	(1.018)	(2.179)	(2.538)

Table 12: Stage 1 Regression Results for Teams Categorized Based on Theory

	Team Revenue			
Dependent Variable	1	2	3	4
Adjusted R-Square	0.9454	0.9549	0.9586	0.9823
Number of Observations	458	458	458	458
Time Trend	12.02**	11.63**	11.36**	6.56**
	(10.238)	(10.829)	(8.995)	(7.722)
Inverse of Team Age		93.41	82.98	19.1
		(1.775)	(1.738)	(0.610)
(Inverse of Team Age)^2		-50.16	-35.68	16.43
		(-0.983)	(-0.613)	(0.432)
Inverse of Stadium Age		165.70**	149.32**	65.15**
		(6.121)	(6.276)	(4.085)
(Inverse of Stadium Age)^2		-126.80**	-109.14**	-43.70*
		(-4.481)	(-4.135)	(-2.502)
Lag of Playoffs Dummy Variable			25.10**	17.23**
			(6.072)	(6.335)
Price				0.48**
				(23.413)

NOTE:
t-statistic in parentheses beneath estimate.
Teams are categorized theoretically based on teams with similar type fan bases.
The categories are as follows:
Category 1: NYY, BOS, NYM, LAD, CHC
Category 2: TEX, LAA, BAL, PHI, SEA, SF
Category 3: CWS, HOU, DET, ATL, STL
Category 4: CLE, MIN, MIL, ARI, COL, SD, TOR
Category 5: TB, PIT, FLA, OAK, KC, CIN
These categories are then interacted with time period dummy variables indicating the time period of the three different Collective Bargaining Agreements (CBA).
The coefficients listed are for each category's intercept, win percentage effect, and square root of win percentage above the .506 threshold effect (listed as WPT for win percentage threshold) within the stated time period.
*Signficant at 5% level.
**Significant at 1% level.

Table 12 Continued: Stage 1 Regression Results for Teams Categorized Based on Theory

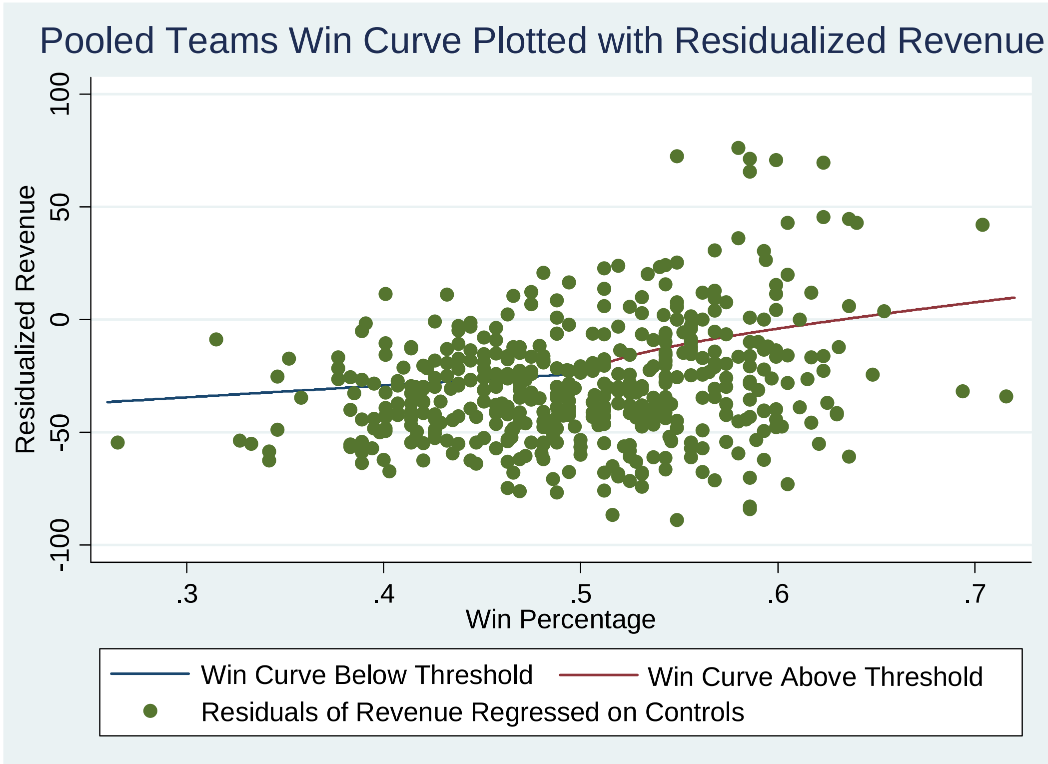
	Team Revenue			
Dependent Variable	1	2	3	4
Category Estimates within First Time Period:				
Category 1 Intercept	-137.92	-136.74	-132.69	-123.94
	(-1.624)	(-1.579)	(-0.913)	(-1.306)
Category 2 Intercept	-48.44	-5.29	28.9	-40.15
	(-0.451)	(-0.050)	(0.252)	(-0.536)
Category 3 Intercept	-51.3	-130.23*	-127.61	-104.2
	(-0.663)	(-2.219)	(-0.963)	(-1.204)
Category 4 Intercept	9.58	-67.07	-79.13	-103.42
	(0.069)	(-1.077)	(-0.675)	(-1.350)
Category 5 Intercept	114.94**	42.23	15.99	-31.49
	(3.968)	(1.526)	(0.211)	(-0.636)
Category 1 Win Percent	500.24**	501.17**	484.53	321.57
	(2.740)	(2.688)	(1.524)	(1.548)
Category 2 Win Percent	276.58	179.05	95.87	126.82
	(1.140)	(0.742)	(0.382)	(0.774)
Category 3 Win Percent	257.83	413.86**	400.78	225.53
	(1.546)	(3.214)	(1.395)	(1.201)
Category 4 Win Percent	105.69	252.65	275.85	230.28
	(0.351)	(1.883)	(1.075)	(1.374)
Category 5 Win Percent	-172.02**	-16.73	45.4	54.04
	(-2.810)	(-0.288)	(0.261)	(0.476)
Category 1 WPT	-53.98	-54.82	-85.02	-25.46
	(-0.533)	(-0.530)	(-0.521)	(-0.239)
Category 2 WPT	-44.63	-73.99	-53.31	-59.73
	(-0.320)	(-0.534)	(-0.402)	(-0.690)
Category 3 WPT	3.12	-112.71	-147.86	-32.86
	(0.036)	(-1.536)	(-1.021)	(-0.347)
Category 4 WPT	137.33	0.82	-30.5	13.52
	(0.935)	(0.009)	(-0.232)	(0.157)
Category 5 WPT	107.07	36.81	-18.18	9.97
	(1.921)	(0.785)	(-0.157)	(0.132)

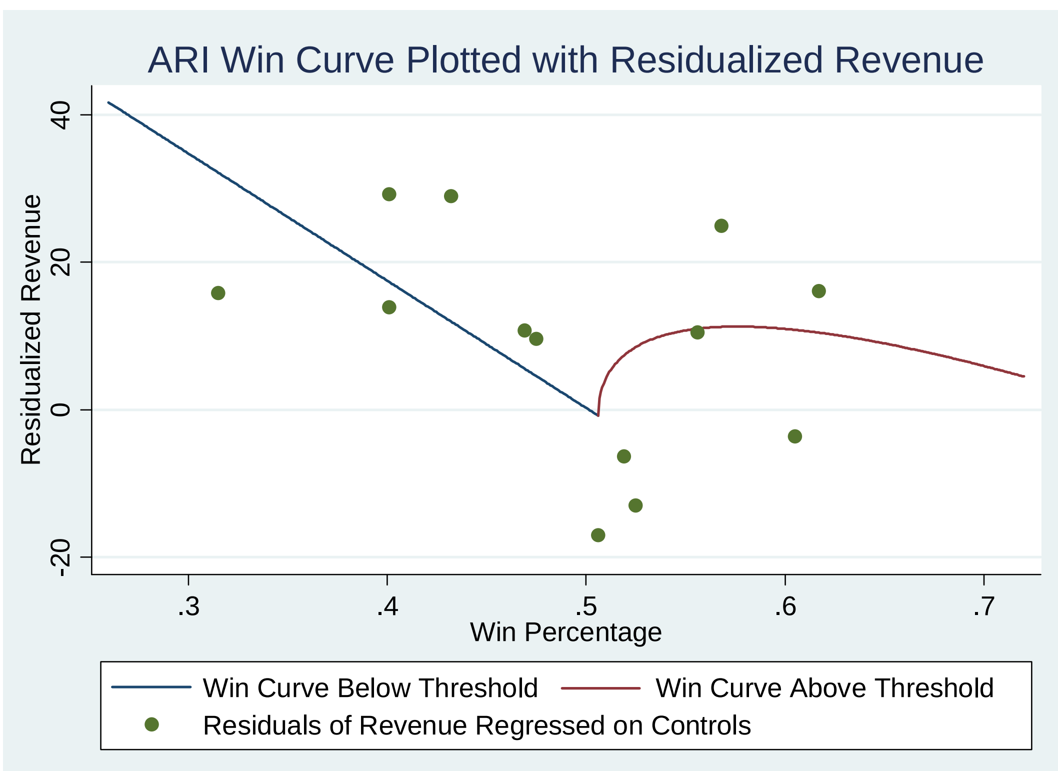
Table 12 Continued: Stage 1 Regression Results for Teams Categorized Based on Theory

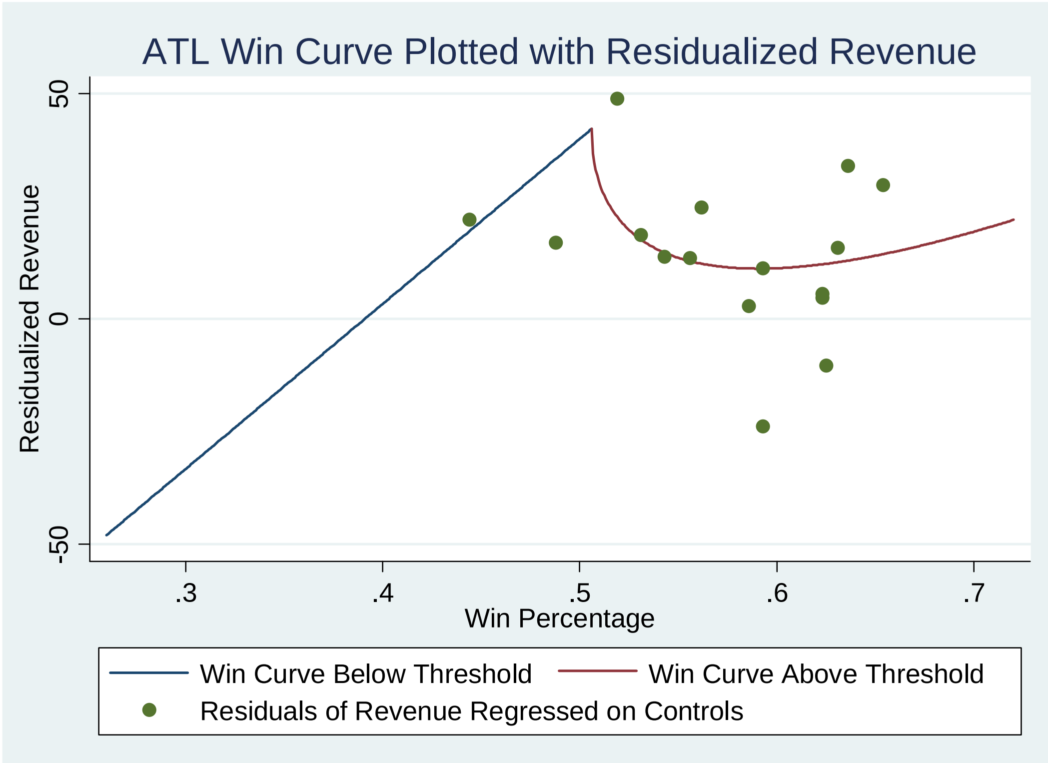
	Team Revenue			
Dependent Variable	1	2	3	4
Category Estimates within Second Time Period:				
Category 1 Intercept	19.21	21.92	23.09	-70.67
	(0.224)	(0.260)	(0.131)	(-0.613)
Category 2 Intercept	213.80**	167.42**	184.77	111.22
	(4.857)	(3.939)	(1.475)	(1.359)
Category 3 Intercept	23.71	-56.66	-24.01	-53.86
	(1.706)	(-1.695)	(-0.365)	(-1.254)
Category 4 Intercept	32.21	15.22	25.72	-13.61
	(1.358)	(0.267)	(0.398)	(-0.322)
Category 5 Intercept	3.62	38.22	42.14	16.26
	(0.157)	(1.683)	(0.638)	(0.377)
Category 1 Win Percent	146.61	148.39	147.51	226.97
	(0.719)	(0.738)	(0.370)	(0.871)
Category 2 Win Percent	-338.98**	-235.81*	-270.33	-195.61
	(-3.518)	(-2.524)	(-0.985)	(-1.091)
Category 3 Win Percent	56.68	202.94**	116.24	107.53
	(1.649)	(2.689)	(0.786)	(1.113)
Category 4 Win Percent	15.54	38.39	17.47	38.3
	(0.274)	(0.305)	(0.120)	(0.404)
Category 5 Win Percent	38.31	-72.79	-73.65	-66.46
	(0.788)	(-1.398)	(-0.487)	(-0.673)
Category 1 WPT	111.8	110.73	65.37	2.42
	(0.853)	(0.854)	(0.303)	(0.017)
Category 2 WPT	233.30**	114.52*	99.49	88.57
	(4.494)	(2.255)	(0.723)	(0.985)
Category 3 WPT	46.54	9.19	30.73	14.77
	(1.318)	(0.174)	(0.321)	(0.236)
Category 4 WPT	-32.3	-56.5	-86.76	-68.17
	(-0.698)	(-0.836)	(-0.945)	(-1.137)
Category 5 WPT	5.67	136.70**	78.67	45.18
	(0.170)	(3.472)	(0.735)	(0.646)

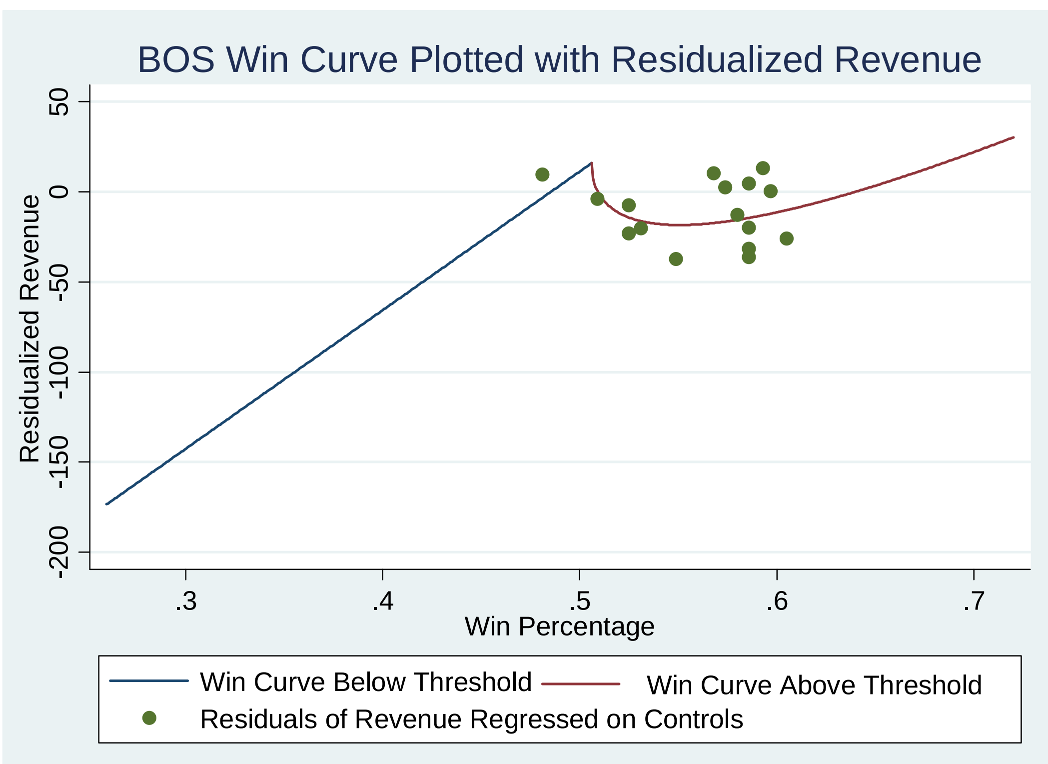
Table 12 Continued: Stage 1 Regression Results for Teams Categorized Based on Theory

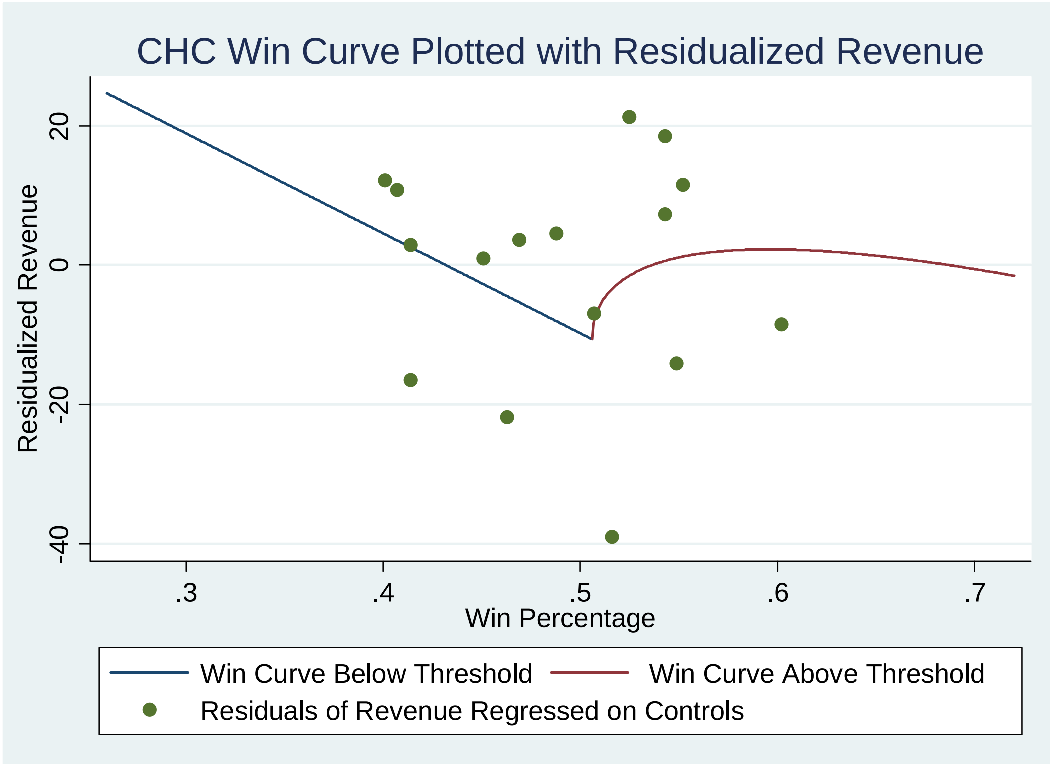
Dependent Variable	Team Revenue 1	2	3	4
Category Estimates within Third Time Period:				
Category 1 Intercept	31.11	-0.4	72.05	-19.89
	(0.091)	(-0.002)	(0.291)	(-0.123)
Category 2 Intercept	-29.75	-26.73	9.71	-72.86
	(-0.622)	(-0.564)	(0.086)	(-0.983)
Category 3 Intercept	197.28	302.04**	416.23	230.2
	(1.584)	(3.005)	(1.532)	(1.296)
Category 4 Intercept	-38.26	-77.03*	19.66	31.27
	(-0.808)	(-2.146)	(0.178)	(0.433)
Category 5 Intercept	-25.35	-14.24	6.11	63.05
	(-0.568)	(-0.364)	(0.067)	(1.061)
Category 1 Win Percent	78.46	132.98	-33.21	74.65
	(0.112)	(0.238)	(-0.064)	(0.219)
Category 2 Win Percent	132.04	135.93	58.91	205.81
	(1.317)	(1.350)	(0.229)	(1.222)
Category 3 Win Percent	-343.59	-569.29**	-817.72	-479.27
	(-1.260)	(-2.602)	(-1.401)	(-1.256)
Category 4 Win Percent	98.62	188.95**	-34.16	-80.86
	(1.067)	(2.767)	(-0.144)	(-0.522)
Category 5 Win Percent	42.55	22.62	-19.45	-175.76
	(0.434)	(0.264)	(-0.096)	(-1.327)
Category 1 WPT	265.24	254.05	293.11	144.34
	(1.112)	(1.101)	(1.349)	(1.016)
Category 2 WPT	22.3	-3.6	-2.86	-75.46
	(0.400)	(-0.063)	(-0.019)	(-0.750)
Category 3 WPT	119.29	190.25*	300.73	185.02
	(1.015)	(2.161)	(1.198)	(1.128)
Category 4 WPT	11.12	-47.6	74.1	78.38
	(0.214)	(-0.924)	(0.589)	(0.954)
Category 5 WPT	-27.4	-3.64	17.74	132.6
	(-0.416)	(-0.063)	(0.137)	(1.568)

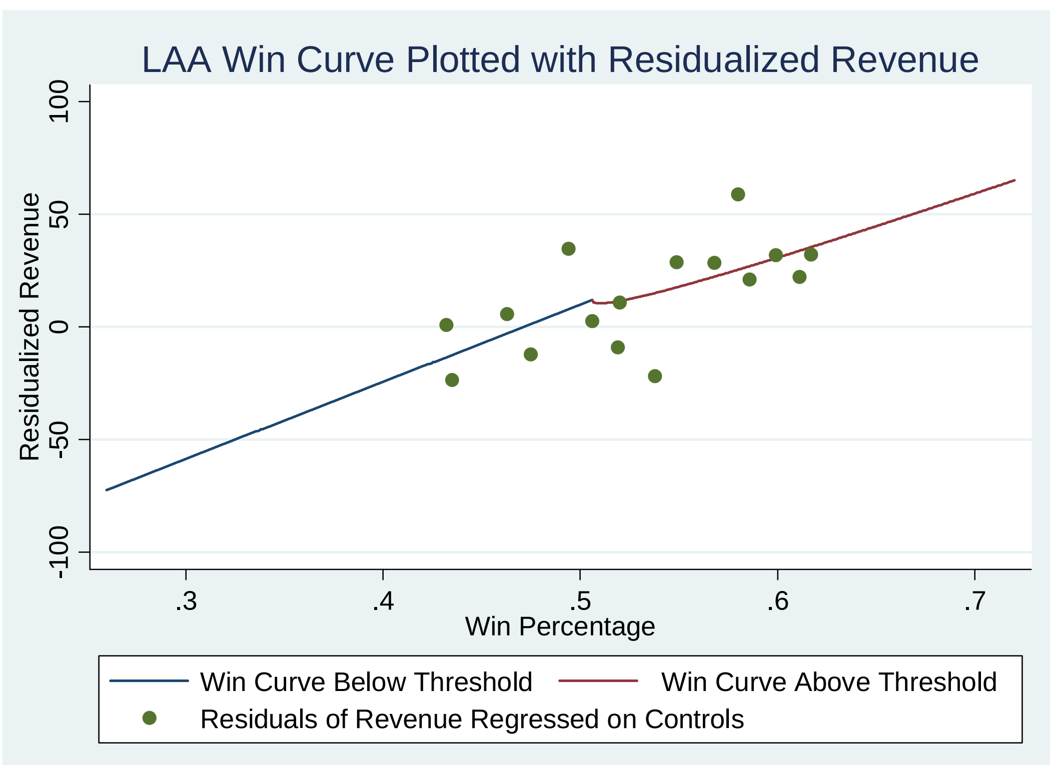

Pooled Teams Win Curve Plotted with Residualized Revenue
Residualized Revenue
100
50
0
-50
-100
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls

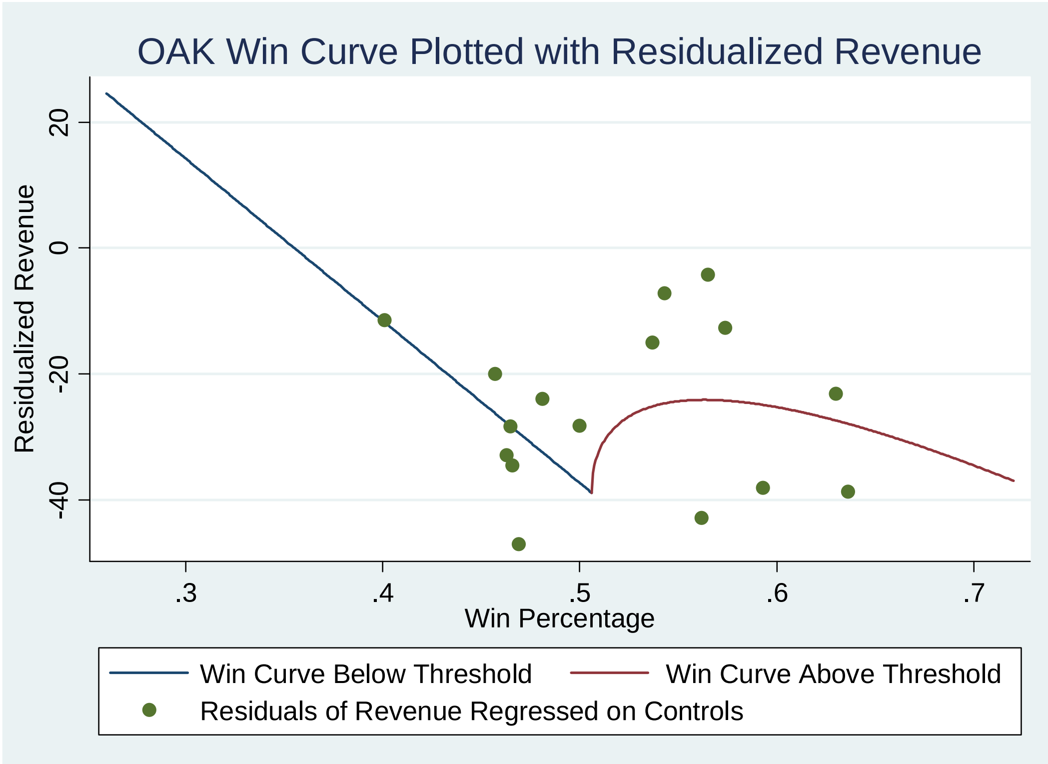

ARI Win Curve Plotted with Residualized Revenue
Residualized Revenue
40
20
0
-20
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls

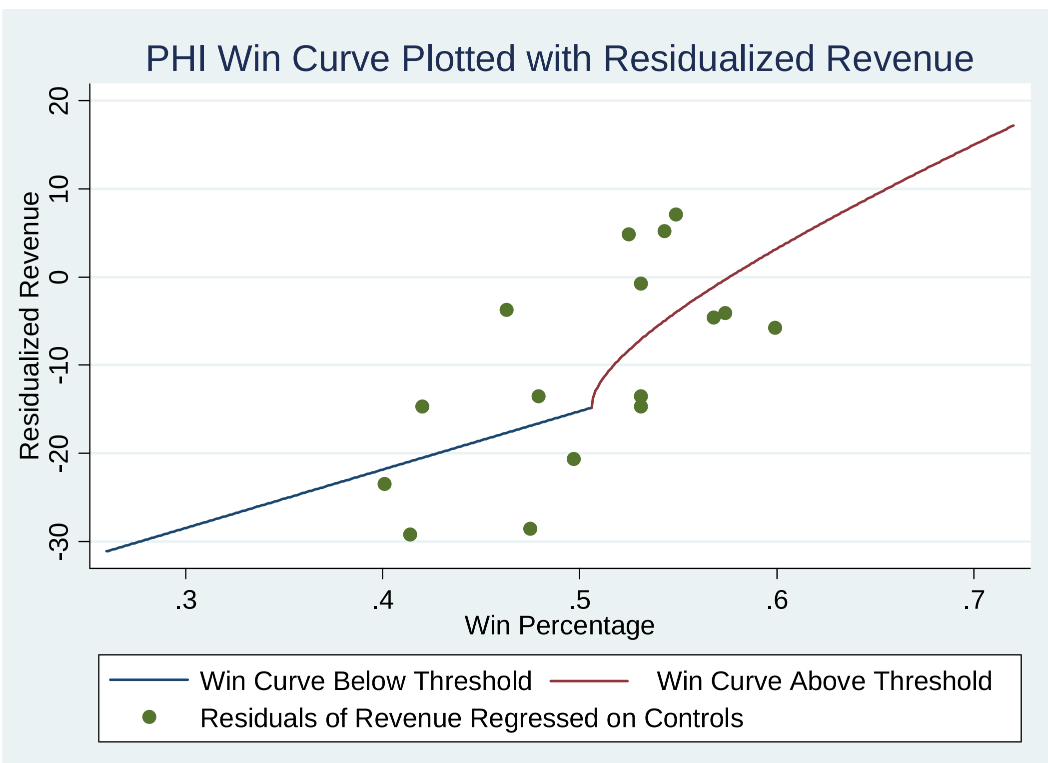

ATL Win Curve Plotted with Residualized Revenue
Residualized Revenue
50
0
-50
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls

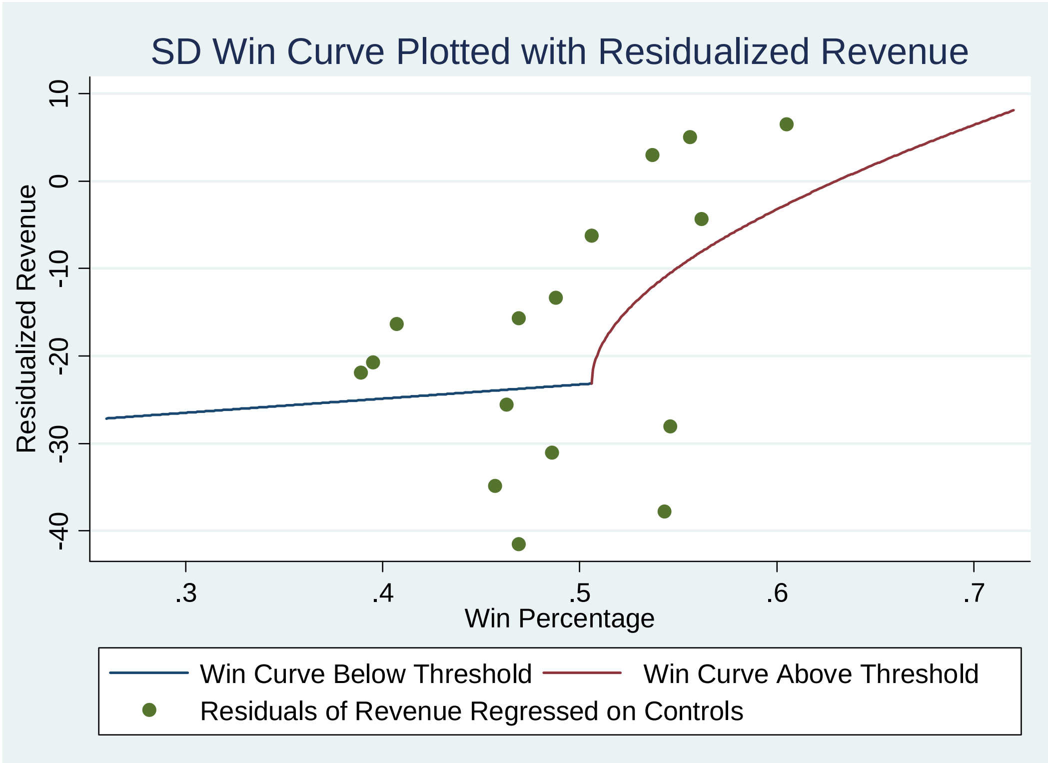

BOS Win Curve Plotted with Residualized Revenue
Residualized Revenue
50
0
-50
-100
-150
-200
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls

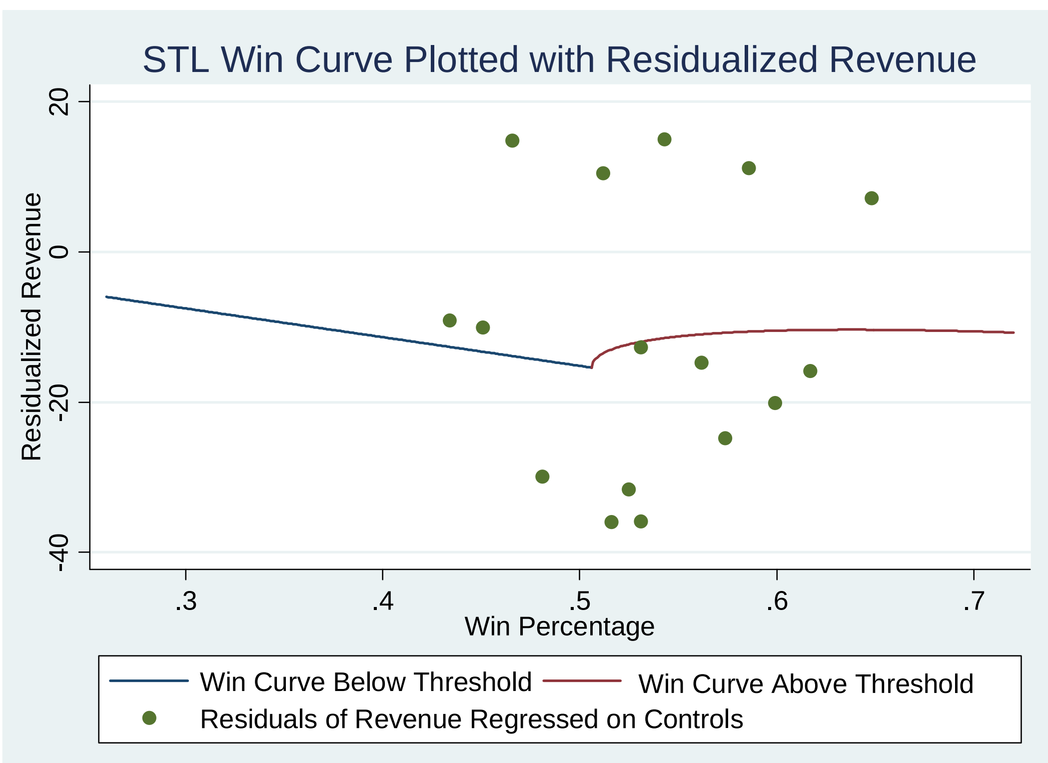

CHC Win Curve Plotted with Residualized Revenue
Residualized Revenue
20
0
-20
-40
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls

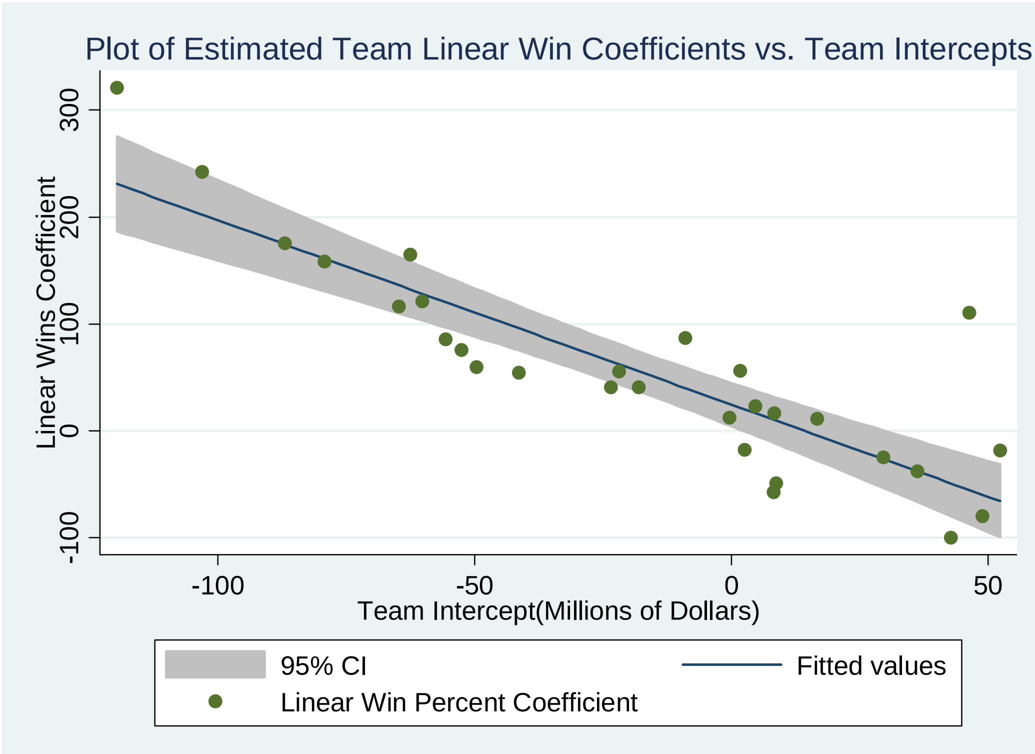

LAA Win Curve Plotted with Residualized Revenue
Residualized Revenue
100
50
0
-50
-100
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls

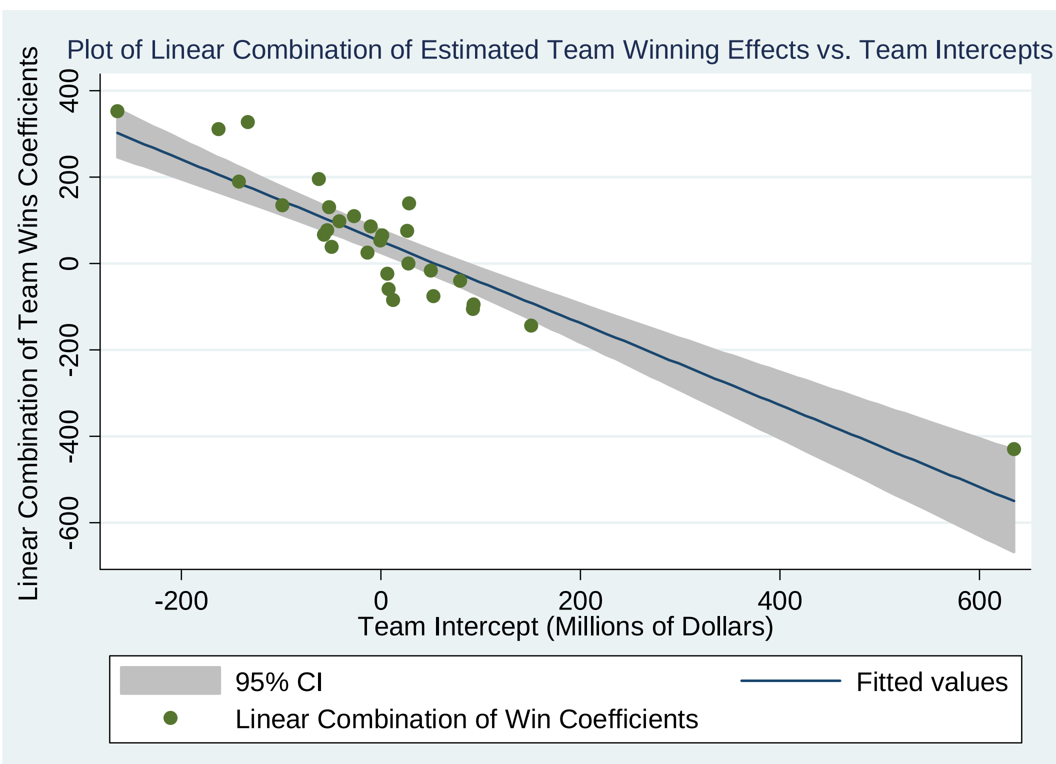

OAK Win Curve Plotted with Residualized Revenue
Residualized Revenue
20
0
-20
-40
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls


PHI Win Curve Plotted with Residualized Revenue
Residualized Revenue
20
10
0
-10
-20
-30
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls


SD Win Curve Plotted with Residualized Revenue
Residualized Revenue
10
0
-10
-20
-30
-40
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls


STL Win Curve Plotted with Residualized Revenue
Residualized Revenue
20
0
-20
-40
.3
.4
.5
.6
.7
Win Percentage
Win Curve Below Threshold
Win Curve Above Threshold
Residuals of Revenue Regressed on Controls


Plot of Estimated Team Linear Win Coefficients vs. Team Intercepts
Linear Wins Coefficient
-100
0
100
200
300
-100
-50
0
50
Team Intercept(Millions of Dollars)
95% CI
Fitted values
Linear Win Percent Coefficient


Plot of Linear Combination of Estimated Team Winning Effects vs. Team Intercepts
Linear Combination of Team Wins Coefficients
-600
-400
-200
0
200
400
-200
0
200
400
600
Team Intercept (Millions of Dollars)
95% CI
Fitted values
Linear Combination of Win Coefficients

References

August, Steven W. Personal interview. Aug. 2011.

Barbarisi, Daniel. Telephone interview. 23 June 2011.

Berri, David J., and Martin D. Schmidt. "What Takes Them out to the Ballgame?" *Journal of Sports Economics* Vol. 7.No. 2 (2006): 222-33.

Blass, Asher A. "Does the Baseball Labor Market Contradict the Human Capital Model of Investment?" *Review of Economics and Statistics* Vol. 74 (1992): 261-68.

Brown, Maury. "MLB's Luxury Tax, Revenue-Sharing, the Yankees, and the Randy Levine Story." *Biz of Baseball*. 7 Apr. 2010. Web. <http://bizofbaseball.com/index.php?option=com_content&view=article&id=4272:mlbs luxury-tax-revenue-sharing-the-yankees-and-the-randy-levine story&catid=26:editorials&Itemid=39>.

Brown, Maury. "The Upcoming CBA and the Battles Within It (Part 2) - Revenue Sharing." *The Hardball Times*. 1 May 2006. Web. 26 Nov. 2011. <http://www.hardballtimes.com/main/article/the-upcoming-cba-and-the-battles-within-it part-2-revenue-sharing/>.

Bruggink, Thomas H., and David R. Rose, Jr. "Financial Restraint in the Free Agent Labor Market for Major League Baseball: Players Look at Strike Three." *Southern Economic Journal* Vol. 56.No. 4 (1990): 1029-043.

Burger, John D., and Stephen K. Walters. "Market Size, Pay, and Performance: A General Model and Application to Major League Baseball." *Journal of Sports Economics* Vol. 4.No. 2 (2003): 108-25.

Burger, John D., and Stephen Walters. "The Existence and Persistence of a Winner's Curse: New Evidence from the (Baseball) Field." *Paper Presented at the 2006 Western Economics Association Meetings, San Diego, CA* (2006).

Cashman, Brian, and Michael Fishman. Personal interview. Dec. 2011.

Cassing, James, and Richard W. Douglas. "Implications of the Auction Mechanism Is Baseball's Free Agent Draft." *Southern Economic Journal* Vol. 47.No. 1 (1980): 110-21.

Clapp, Christopher M., and John K. Hakes. "How Long a Honeymoon? The Effect of New Stadiums on Attendance in Major League Baseball." *Journal of Sports Economics* 6.3 (2005): 237-63.

Duquette, Dan. Telephone interview. 2 Jan. 2012.

"Dynamic Pricing Drives Significant Returns During 2011 Major League Baseball Season." *EON: Enhanced Online News*. 25 Oct. 2011. Web. 26 Nov. 2011. <http://eon.businesswire.com/news/eon/20111025006511/en/dynamic-pricing/mlb tickets/new-york-mets>.

Gennaro, Vince. *Diamond Dollars: The Economics of Winning in Baseball*. Hingham, MA: Maple Street, 2007.

Krautmann, Anthony C. "What's Wrong with Scully-estimates of a Player's Marginal Revenue Product." *Economic Inquiry* Vol. 37.No. 2 (1999): 369-81.

Raimondo, Henry J. "Free Agents' Impact on the Labor Market for Baseball Players." *Journal of Labor Research* 4.2 (1983): 183-93.

Rosen, Sherwin. "The Economics of Superstars." *American Economic Review* Vol. 71.No. 5 (1981): 845-58.

Silver, Nate. "Is Alex Rodriguez Overpaid." *Baseball between the Numbers: Why Everything You Know about the Game Is Wrong*. By James Click and Jonah Keri. New York: Basic, 2006. 174-98.

Sommers, Paul M., and Noel Quinton. "Pay and Performance in Baseball: The Case of the First Family of Free Agents." *Journal of Human Resources* Vol. 17.No. 3 (1982): 426-36.

Stock, James H., and Mark W. Watson. "Generalized Least Squares." *Introduction to Econometrics*. Boston: Pearson/Addison Wesley, 2007. 724. Print.

"Team Marketing Report 2011 Fan Cost Index." Team Marketing Report, Apr. 2011. Web. <https://www.teammarketing.com/public/files/2011_mlb_fci.pdf>.

William, Gould. "FAQ: Clarification on Analytic Weights with Linear Regression." *Stata: Data Analysis and Statistical Software*. StatCorp, Jan. 1999. Web. 24 Apr. 2012. <http://www.stata.com/support/faqs/stat/crc36.html>.

Zimbalist, Andrew S. "Competitive Balance in Major League Baseball." *Milken Institute Review* Vol. 3 (2001): 54-64.

Zimbalist, Andrew S. "Salaries and Performance: Beyond the Scully Model." *Diamonds Are Forever: The Business of Baseball*. By P. M. Sommers. Washington, D.C.: Brookings Institution, 1992. 109-33.