Can Money Buy You Playoff Spots and Championships in Major League Baseball?

Jay R. Schaffer
University of Northern Colorado

The New York Yankees have long been thought of as the 800 lbs gorilla in the room when it comes to baseball payrolls. For years the Yankees have out spent every other team in major league baseball to "buy championships". Sometimes it has paid off (championships in 2000 and 2009); other times it has not.

How much have they spent? In 2009, the Yankees offered contracts to C.C. Sabathia (7 years, $161 million), A.J. Burnett (5 years, $82.5 million), and Mark Teixeira (8 years, $180 million) on top of the contracts already offered to Alex Rodriguez (10 years, $275 million) and Derek Jeter (10 years, $189 million). These were staggering amounts even by the New York Yankees standard.

The question to be asked by the rest of the teams in major league baseball is "does money buy playoff spots and championships?" This research argues that big dollar team payrolls do give an unfair advantage to some teams in major league baseball. The data from 2000-2009 seems to support that claim.

Methodology

Opening day team payrolls and the number of wins a team obtained in a season were collected for 2000-2009 from http://www.stevetheump.com/Payrolls.htm and http://www.baseball-reference.com/. A simple linear regression model from Montgomery and Peck (1992), shown in Equation 1, was fit to the data for each year.

$$Y = \beta_0 + \beta_1 X + \varepsilon \qquad (1)$$

The results are shown in Figures 1-10 below. Teams that made the playoffs are denoted with white diamonds, while teams that did not make the playoffs are denoted with black diamonds. The World Series champion is denoted by a white triangle.

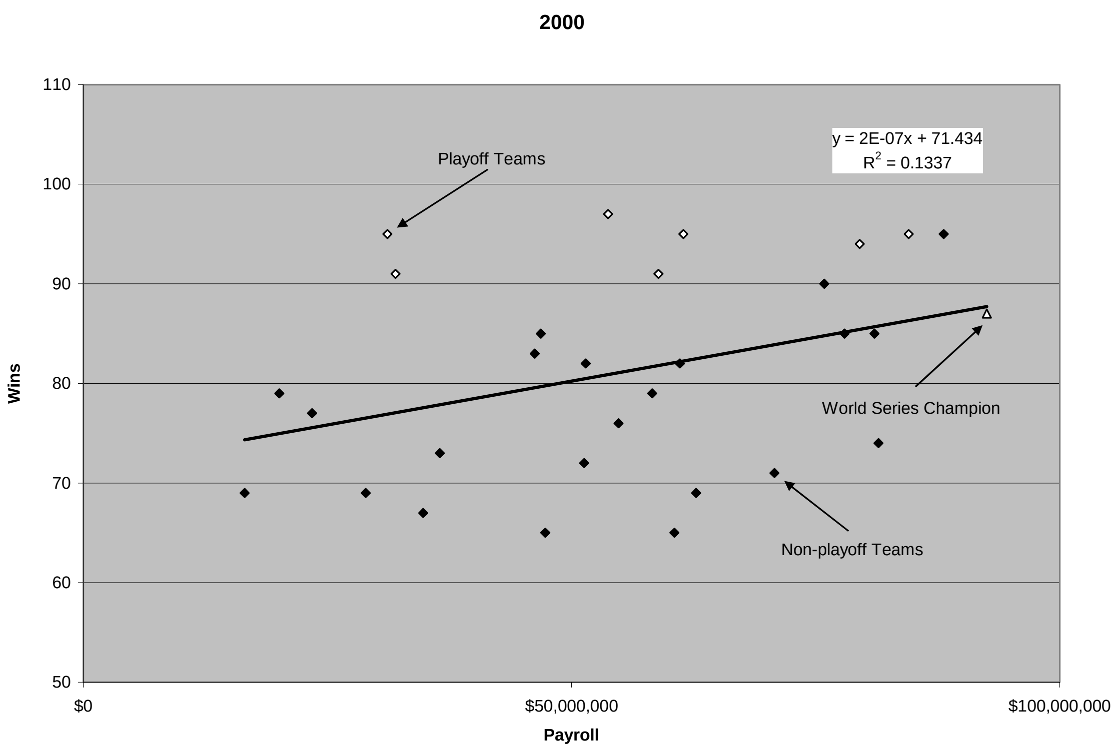

Figure 1. 2000 Season

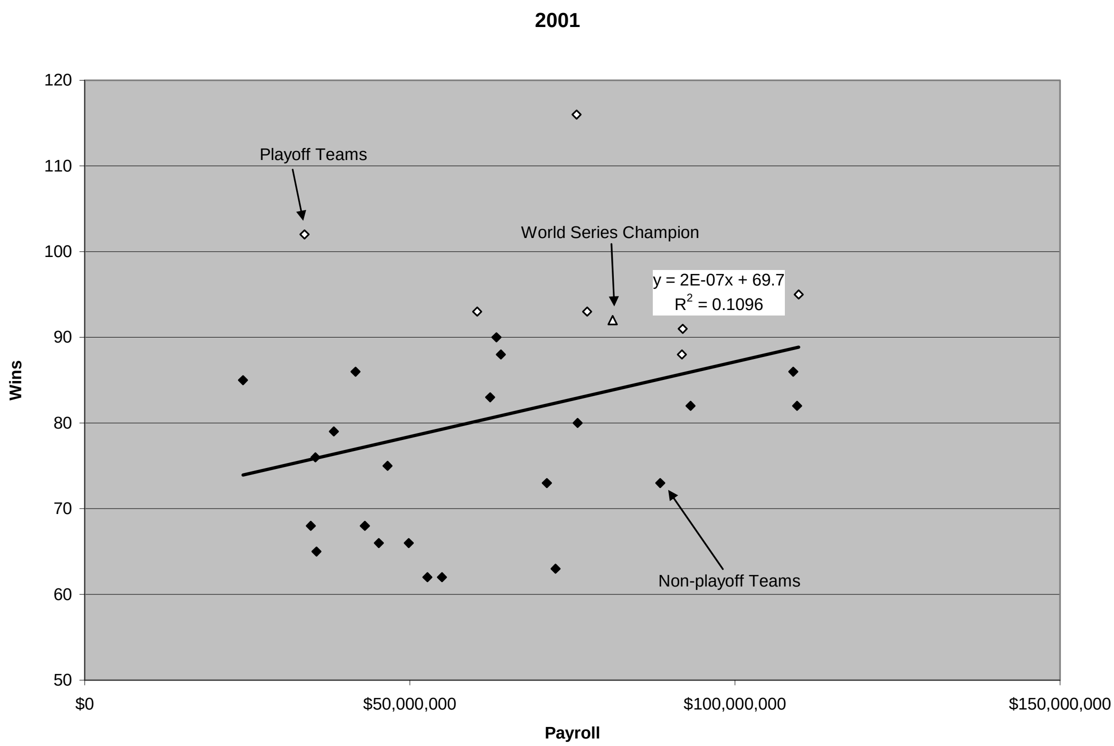

Figure 2. 2001 Season

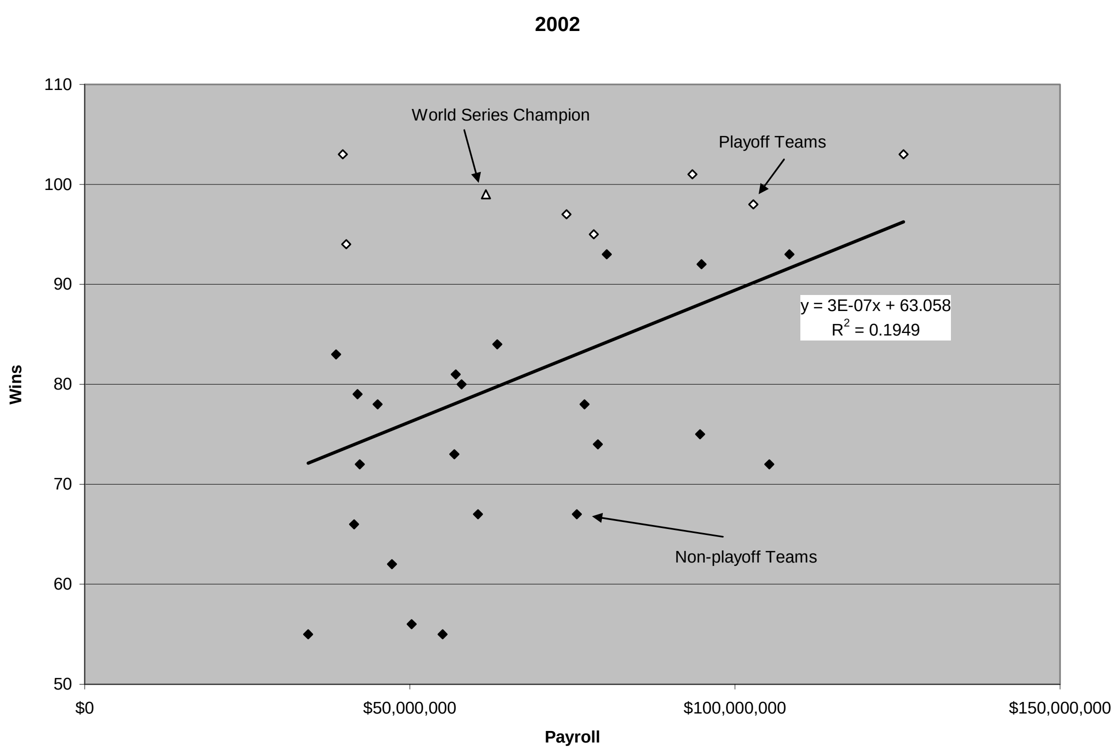

Figure 3. 2002 Season

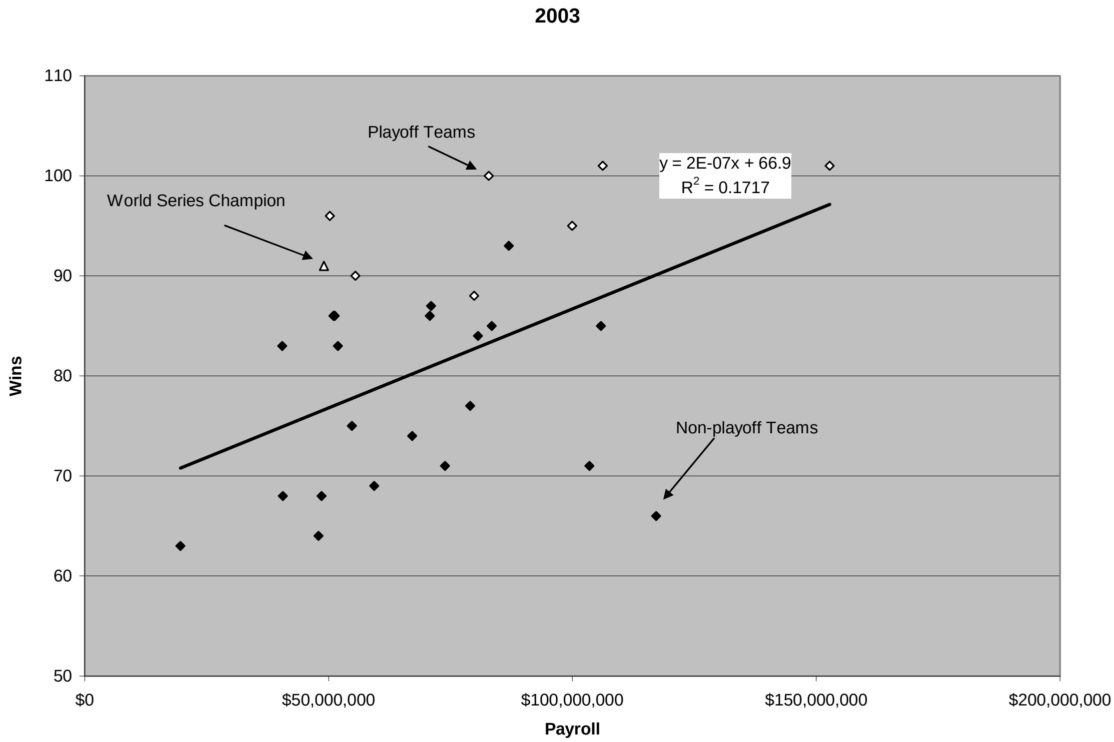

Figure 4. 2003 Season

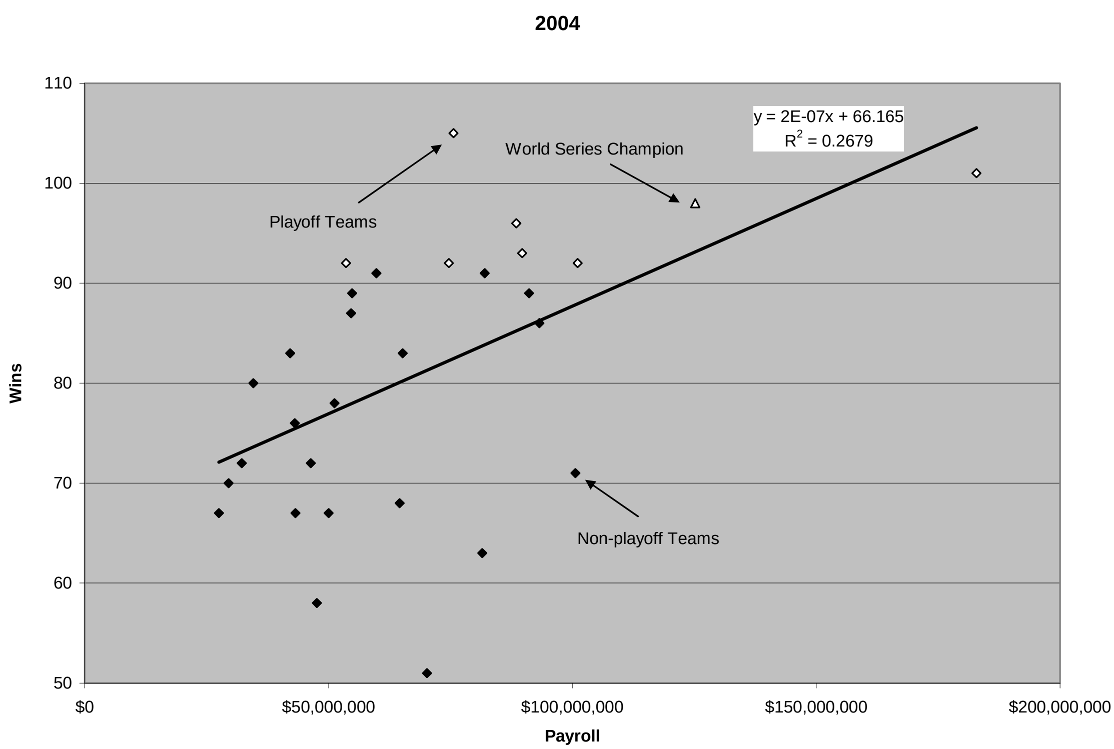

Figure 5. 2004 Season

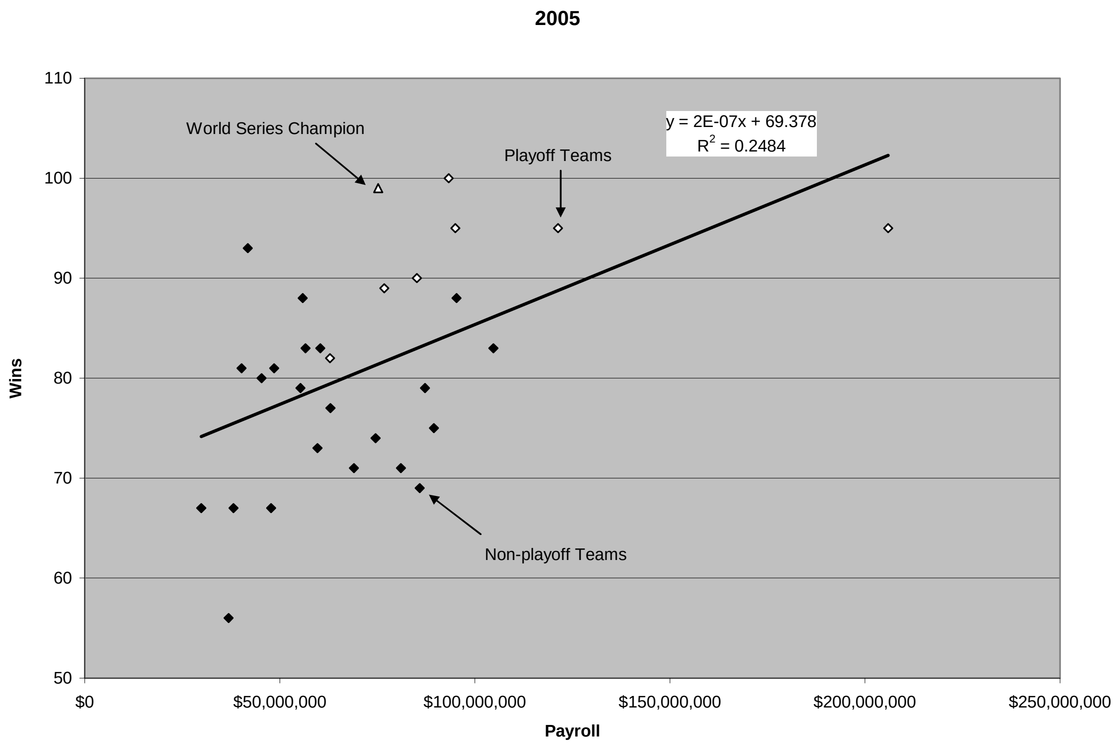

Figure 6. 2005 Season

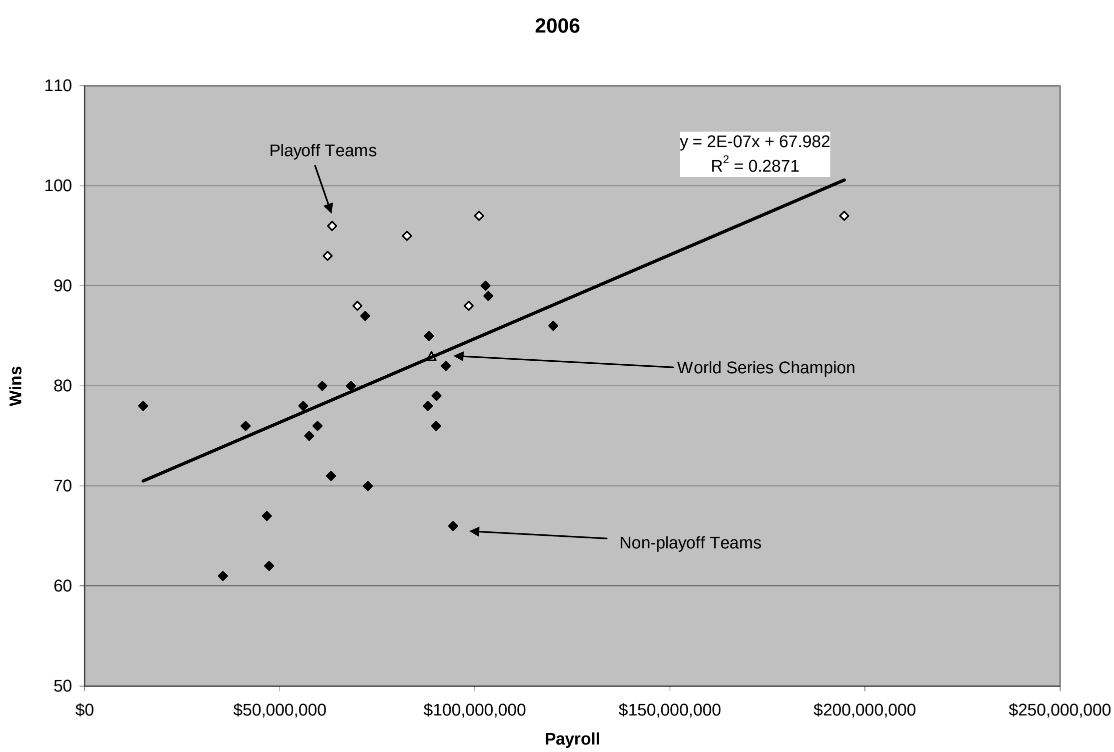

Figure 7. 2006 Season

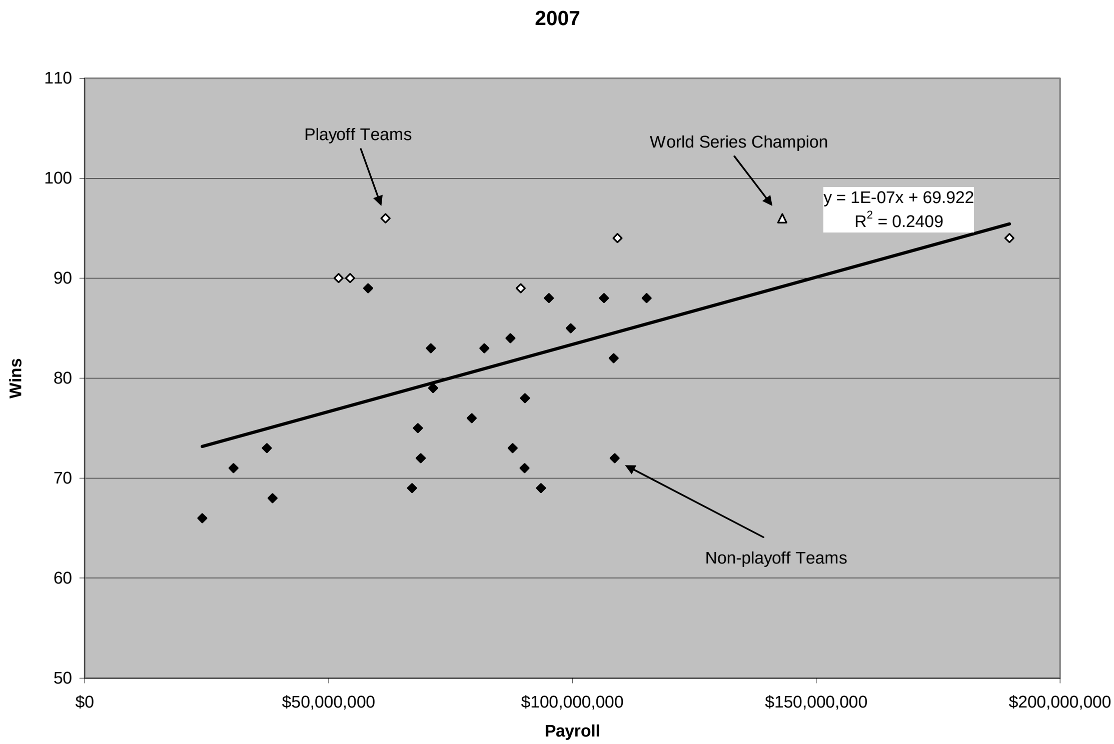

Figure 8. 2007 Season

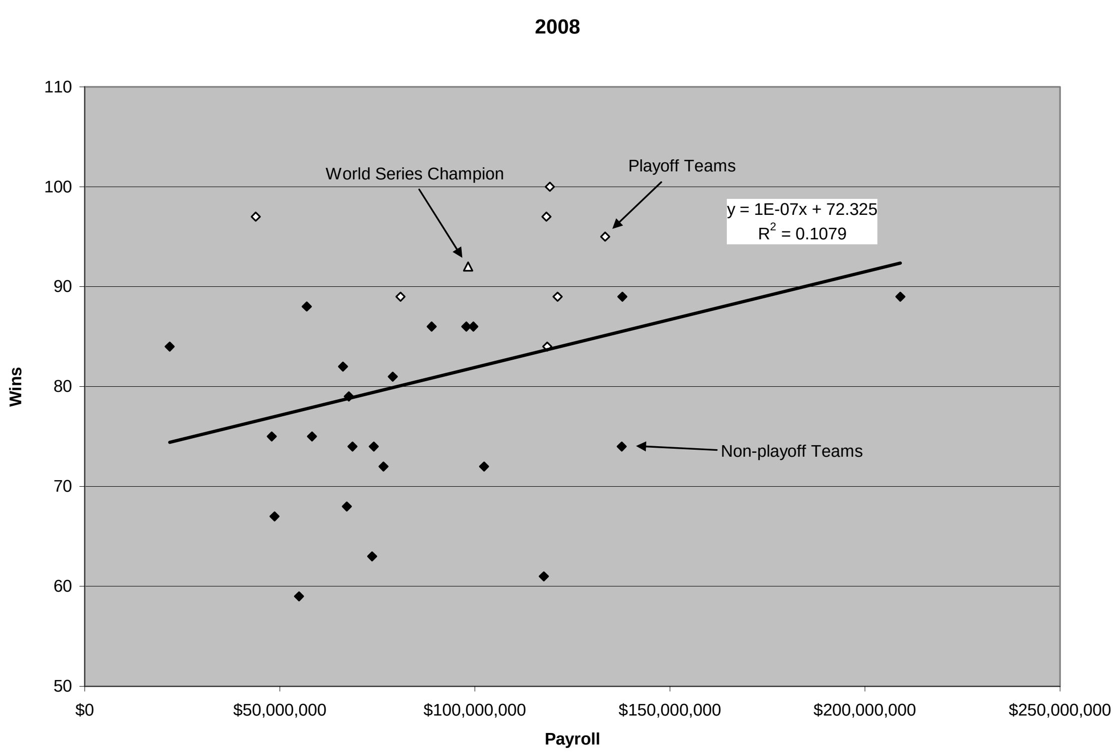

Figure 9. 2008 Season

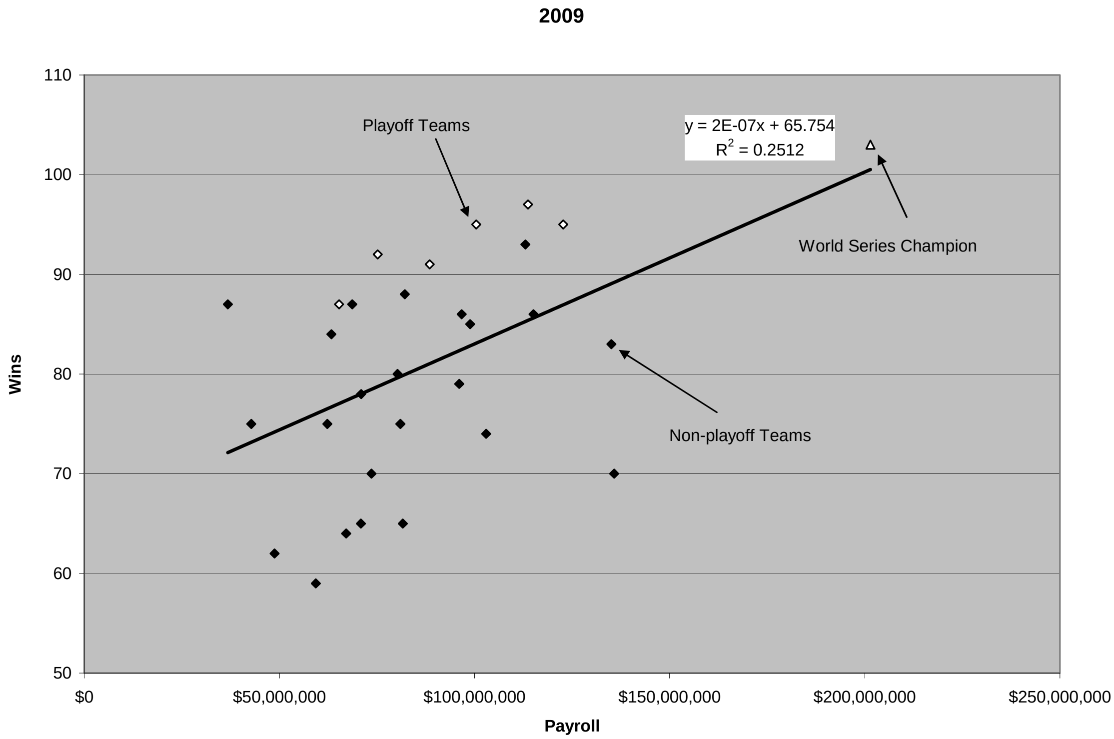

Figure 10. 2009 Season

Table 1. Coefficients of Determination, Regression Parameter Estimates, and p-values for H_0: $\beta_1 = 0$.

Season	r^2	b_0	b_1	p
2000	0.1337	71.43	1.76×10^{-7}	0.0469
2001	0.1096	69.69	1.74×10^{-7}	0.0740
2002	0.1949	63.06	2.63×10^{-7}	0.0146
2003	0.1717	66.90	1.98×10^{-7}	0.0228
2004	0.2679	66.16	2.15×10^{-7}	0.0034
2005	0.2484	69.38	1.60×10^{-7}	0.0051
2006	0.2871	67.98	1.67×10^{-7}	0.0023
2007	0.2409	69.92	1.34×10^{-7}	0.0059
2008	0.0803	72.98	7.99×10^{-8}	0.1292
2009	0.2512	65.75	1.72×10^{-7}	0.0048

Table1 contains the coefficients of determination, regression parameter estimates, and *p*-values for H_0: $\beta_1 = 0$ obtained from Equation 1 for each season. It should be noted that a statistically significant slope (i.e. relationship between team payroll and wins) was found in eight of the ten seasons using $\alpha = 0.05$.

In addition, it should be noted that the average r^2 for the 10 seasons is 0.1986 meaning that nearly 20% of the variation in wins is being explained by team payroll. It also appears that in recent years (2004-2009), the relationship between team payroll and wins has grown stronger (e.g. 2006 season $r^2 = 0.2871$, 2009 season $r^2 = 0.2512$). While these r^2 values appear small, low r2 values have been common in previous baseball studies. Schaffer and Heiny (2006) analyzed major league baseball data from the 2003 season. They were attempting to measure the effect of elevation on slugging percentage. They used slugging percentage as a dependent variable and elevation, ballpark, and ball player effects as independent variables. The r^2 for their model was 0.21. Hofacker (1988) analyzed major league baseball data from the 1982 season. He was attempting to measure a team's offensive ability independent of opponent and ballpark. He used runs scored as the dependent variable and opponent, park, league and home vs. away as independent variables. The r^2 for his model was 0.267 and he had the following

Table 2. Number of Playoff Teams from (max-Q_3), (Q_3-Q_2), (Q_2-Q_1), (Q_1-min) for each season.

Season	Max-Q_3	Q_3-Q_2	Q_2-Q_1	Q_1-Min
2000	3*	2	1	2
2001	4*	2	1	1
2002	3	3*	0	2
2003	3	2	1	2*
2004	5*	2	1	0
2005	4	3*	1	0
2006	3	2*	3	0
2007	4*	1	0	3
2008	5	2*	0	1
2009	4*	2	1	1
Totals	38	21	9	12
Percentage	0.4750	0.2625	0.1125	0.1500

Table 3. Number of Playoff Teams with the Highest Team Payroll, 2nd Highest Team Payroll, etc. Within Each Division

Playoff Team Payroll Rank within Divsion	1	2	3	4	5
Observed	30	20	15	8	7

comments. "While it is true that researchers in some fields might scoff at such a low r^2, perhaps the better way to think about the current result is that it offers insight into just how stochastic baseball must be. Such considerations necessarily imply that the analysis presented be considered exploratory." Another explanation might be that additional variability in the model is lost when players get injured or teams underperform during the season independent of team payroll. The important item to note is that these studies had several independent variables in their models while the current study has but one independent variable, yet explains just as much, and sometimes more, variability.

Additional evidence seems to indicate that team payroll influences the number of wins. For each season, the team payroll quartiles (Q1, Q2, Q3) were obtained and the number of teams that made the playoffs from (max-Q3), (Q3-Q2), (Q2-Q1), and (Q1-min) were determined. Table 2 below displays the quartile data. The World Series champion is denoted by (*).

From Table 2, it should be noted that 38 of the 80 playoff teams came from the (max-Q_3) quartile while only 12 came from the (Q_1-min) quartile. In addition, 9 of the 10 World Series champions came from the (max-Q_3) or (Q_3-Q_2) quartiles. The lone exception was the 2003 World Series champion Florida Marlins.

A Chi-squared Goodness of Fit Test from Conover (1999) was conducted to examine the question of equal dispersion of playoff teams from (max-Q_3), (Q_3-Q_2), (Q_2-Q_1), and (Q_1-min). If there was an equal dispersion of playoff teams, one would expect to find 20 playoff teams from each quartile. Equation 2 below displays the Chi-squared Goodness of Fit statistic and corresponding calculations where O_i is the observed number of playoff teams from (max-Q_3), (Q_3-Q_2), (Q_2-Q_1), and (Q_1-min), E_i is the expected number of playoff teams, and N is the total number of playoff teams.

$$\chi^2 = \sum_{i=1}^{4} \frac{O_i^2}{E_i} - N = \left(\frac{38^2}{20} + \frac{21^2}{20} + \frac{9^2}{20} + \frac{12^2}{20} \right) - 80 = 25.5 \quad (2)$$

The corresponding *p*-value of the Chi-squared Goodness of Fit statistic is less than 0.001 indicating that too many teams come from above the median and too few from below the median.

Investigating the imbalance further, team payrolls of playoff teams were ranked within baseball divisions. Table 3 shows how many teams made the playoffs with the highest payroll, 2nd highest payroll, etc within divisions.

A Chi-squared Goodness of Fit Test from Conover (1999) was conducted to examine the question of equal dispersion of playoff teams within team payroll rank. If there was an equal dispersion of playoff teams, one would expect to find 16 playoff teams from each level. Equation 3 below displays the Chi-squared Goodness of Fit statistic and corresponding calculations where O_i is the observed number of playoff teams from each division rank, E_i is the expected number of playoff teams, and N is the total number of playoff teams.

$$\chi^2 = \sum_{i=1}^{4} \frac{O_i^2}{E_i} - N = \left(\frac{30^2}{16} + \frac{20^2}{16} + \frac{15^2}{16} + \frac{8^2}{16} + \frac{7^2}{16} \right) - 80 = 22.4 \qquad (3)$$

The corresponding p-value of the Chi-squared Goodness of Fit statistic is less than 0.001 indicating that too many teams come from teams with the highest team payroll within a division and too few from the lowest team payroll with a division. It would appear if one was to bet on playoff spots, simply bet on the team with highest payroll within the division. An example of this is the 2003 Minnesota Twins. They had the 18th highest team payroll overall, but had the highest team payroll within the American League Central Division and made the playoffs.

Some would argue that the luxury tax (or competitive balance tax) was supposed to level the playing field. However, according to http://www.stevetheump.com/luxury_tax.htm, the New York Yankees are essentially the only team to ever exceed the luxury tax salary cap. For example, in 2008, the Yankees were charged a 40% penalty for exceeding the team cap of $155 million dollars. The Yankees final payroll of the season was $222.2 million and had to pay $26.9 million in tax that was distributed to the other major league teams. However, when $26.9 million is divided up by the 29 remaining teams, less than $1 million is being added to their respective payrolls. $1 million barely covers two players making the league minimum ($400,000). According to the regression models discussed above, adding $1 million to a team's payroll will not generate many additional wins or playoff spots. So it would appear that the luxury tax is a failure.

Others would argue that the current revenue sharing agreement was supposed to create a better competitive balance among the 30 teams. According to Ray (2007), "In 1997, major league baseball created a new revenue sharing system that requires successful teams to pay millions of dollars every year to unsuccessful teams." Unfortunately Ray states, "The revenue sharing agreement doesn't require recipients to spend the "shared" revenue on actual ballplayers. All that is required by teams is that they use the money "to improve the product on the field." That vague requirement, however, has not been enforced by the League. In reality, the money can go anywhere. It can even go into the owner's pockets." Ray adds "From 2002 through 2006, Tampa Bay took in an average of $32 million per year in revenue sharing money. During that same period, the Rays had an average payroll of just $27 million, which was the lowest in baseball. They also had the worst five year record on the field, winning an average of just 70 games per season. Yet the team turned an average profit of more than $20 million during those years." So it would appear the revenue sharing agreement needs to be revisited.

Conclusion

It would appear from the evidence presented that team payrolls unduly influences the number of wins obtained in any given season, playoff spots obtained, and championships won. Money can improve a team's chance of obtaining a playoff spot and shot at a championship. The data seems to support that a more equitable system for team payrolls must be put in place. Either a salary cap needs to be imposed or a minimum team payroll must be enforced by Major League Baseball under the current revenue sharing rules. The inequities in the data are obvious. Big payroll teams like the New York Yankees or Los Angles Angels are over represented in the playoffs while small payroll teams like the Pittsburgh Pirates, Kansas City Royals, or Florida Marlins are not competitive under the current system.

References

Conover, W. J., (1999). *Practical Nonparametric Statistics*. (3rd ed.). New York: John Wiley & Sons, Inc.

Hofacker, Charles F. (1988) Removing Extraneous Effects from Baseball Performance Measures. *Journal of Sport Behavior*, Vol. 11, No. 3., 204-212.

Montgomery, D.C. & Peck, E.A., (1992). *Introduction to Linear Regression Analysis*. (2nd ed). New York: John Wiley & Sons, Inc.
Ray, J.L., (2007). Baseball's Revenue Sharing Problem. *Suite101.com*, *http://baseball.suite101.com/article.cfm/baseballs_revenue_sharing_problem*.
Schaffer, J.R. & Heiny, E. L., (2006). The Effects of Elevation of Slugging Percentage in Major League Baseball. *Chance*, Vol. 19, No. 1, 28-34
www.baseball-reference.com (2009)
www.stevetheump.com/luxury_tax.htm (2009)
www.stevetheump.com/Payrolls.htm

Send correspondence to: Jay R. Schaffer
University of Northern Colorado
Email: Jay.Schaffer@unco.edu