Does Increasing Spending Lead to an Increase in Wins in Major League Baseball

Anthony Brennan

May 8, 2017

One payroll trend that has continued to be used in the sabermetrics era, is that general managers will rapidly increase their team’s payroll when attempting to win more games. This is often done to take an average team and add more talent over the offseason to make them a playoff team the following season. Some teams find success with this strategy. The 2015 Arizona Diamondbacks, for example, increased their payroll by $32.89 million (50% increase) by signing some major free agents and renegotiating contracts, and won 15 more games than they did in 2014. The 2012 Dodgers, following a similar strategy, increased their payroll by $121.16 million and only won 4 more games compared to 2011. This strategy is much easier to practice in the MLB than in the NBA, NFL, and NHL because baseball does not have a salary cap. On the flip side, some teams will shed payroll drastically in an attempt to lose more games (a strategy known as tanking), so they can get a better draft position. This can also occur when the owner isn’t willing to dedicate as much money towards payroll. The 2011 Twins’ payroll was $18.65 million smaller than in 2010 and they lost 31 more games. The point of this paper is to see if making major changes to payroll is still an effective strategy, thus testing to determine if there is a relationship between increasing and decreasing total payroll and change in regular season games won between years.

Since the publishing of the book *Moneyball*, by Michael Lewis, the way Major League Baseball teams (and other sports leagues to an extent) approach building rosters. Lewis followed Billy Beane, general manager of the Oakland Athletics, and the rest of his front office for the entire 2002 season. Being a small market team, the A’s did not have the financial resources to have a large payroll, like the New York Yankees, and struggled to hold on to star players. So Beane and his staff, strapped for cash, took a sabermetric approach to build a more money efficient team. Instead of looking at the more traditional statistics, like home runs and batting average, they used more efficient statistics, like on-base percentage and slugging percentage, to find undervalued players. This strategy was successful and the A’s led the league with 103 wins in 2002. As a result, many teams began to adopt sabermetrics and the moneyball strategy. Other small market teams have found success, with the Cleveland Indians and Tampa Bay Rays both making the World Series on a moneyball payroll. Other big market teams, like the Chicago Cubs, have used sabermetrics to become more efficient. Moneyball and sabermetrics have made the playing field between teams with large and small payrolls a little more even in recent years. In 2016, a low budget Indians team (24th in preseason payroll) finished third in total wins.

Another trend that has become more popular in recent years is the idea of tanking to build from the ground up through the draft and minor leagues, also known as a team’s farm system. The idea is that teams entering the rebuilding phase will trade off most of their talented players for prospects and draft picks, in the hopes that stockpiling young talent will bring them success in the future. With all of the data that is now widely available for talent scouts, teams are able to find rebuild more efficiently. In the past, many teams preferred rebuilding more rapidly by signing free agents, rather than through the farm system. Two teams that went through full rebuilds recently are the Chicago Cubs and the Houston Astros. Both teams released most of their talent for prospects and draft picks. They both suffered through 100 loss seasons in the early 2010’s, while they stock piled young talented players in their farm systems. Over the past couple years these players have risen to the majors, and both teams experienced major increases in wins, the Cubs won 103 games and the World Series in 2016. Other teams, like the Atlanta Braves and Philadelphia Phillies, are currently in the rebuilding process and seem poised for future success. So given these new trends it is necessary to see if these payroll strategies have a net effect on wins.

A couple of studies on the relationship between wins and payroll have been done in the past. Both study the relationship between total payroll on team success, however, whereas this is testing the relationship between change in payroll and change in wins, but they are still relavent. In a study by Shahriar Hasan (2008), he tested to see if the ratio of the average team payroll to the league average payroll had a positive relationship with win percentage between 1992 and 2007. He included a dummy fixed effect for each team to control for the expansion of four new teams and relocation of another during that time period. Hasan found there to be a strongly significant positive relationship and concluded that large payroll is positively correlated with win percentage. Schwartz and Zarrow (2009) come to a similar conclusion in their study. They found a positive relationship between total payroll and win percentage in the regular season between 1977 and 2008. However, they concluded that there is no such relationship in the playoffs. They concluded this is because the results of playoff games are more random and the best regular season teams often don’t win the World Series, the late 90s Yankees dynasty being a historical anomaly.

The data is a panel set for the 2004 (first full season after the publishing of Moneyball) thru the 2015 seasons (with data from 2003 to calculate the change in wins, payroll, and attendance for the 2004 season). Given there are 30 MLB teams and 12 years of data, there will be a total of 360 observations. All of the data is for regular season games only. The data is not a random sample as it includes all of the possible observations for that time frame. There are three major types of data, team wins, payroll, and attendance. The data originates from two different websites. Team wins comes from baseball-reference.com. Thebaseballcube.com is the source for team payroll and team attendance. Each data point is matched up to the year it occurred in and its corresponding team.

There will be seven variables used in the test. The dependent variable will be the change in team wins (*delta\_wins*). This variable is the increase or decrease in total team wins for the current season compared to the previous season, for example the Cleveland Indians won 96 games (out of a 162-game season) in 2007 and won 78 games in 2006, so the change in team wins for 2007 is +18. The first two independent variables will be team payroll (*payroll*) and change in team payroll (*delta\_payroll*), both measured in millions of dollars. *delta\_payroll* is to test the relationship between the change in a team’s payroll and change in wins. *payroll* is the total payroll for each team for the season. It is included to see if there is a difference in the relationship between the change in a team’s payroll and change in wins is different for teams with already large and small payrolls. The next two independent variables are team attendance (*attendance*) and change in attendance (*delta\_attendance*) measured in thousands of fans. Team attendance is related payroll because a large chunk of the amount of money a team is able to spend on payroll is dependent on ticket sales, so this will be used as a control on revenue. The next variable is a control for team wins from the previous year (*previous\_wins*). I included this variable to take into account that it is much harder for a team to increase their total wins if they won 100 games the previous season, than it is if they won 60 the previous season. The final variable is a control for the change in average pay role of all 30 teams (*delta\_avgPay*). Each year the average payroll changes, often positively, based on T.V. deals, ticket prices, apparel sales, etc. This variable will be used to test if there is an effect if teams don’t keep up with the league average.

There will be two econometric models tested, both using delta\_wins as the dependent variable. The first model does not take into account attendance, while the second does.

Model 1:

delta\_wins = B0 + B1\*delta\_payroll + B2\*payroll + B3\*previous\_wins + B4\*delta\_avgPay

Model 2:

delta\_wins = B0 + B1\*delta\_payroll + B2\*payroll + B3\*delta\_attendance + B4\*attendance + B5\* previous\_wins + B6\*delta\_avgPay

The first model leaves out the attendance controls, but keep the controls for the previous season’s wins and change in average league payroll. The second model includes all independent variables. The purpose of controlling for attendance in one and not the other is to see if revenue is playing a factor in the team’s decision to change its payroll. If there is a major drop in attendance, that could force teams to cut back on payroll so that they don’t lose money. So the first model tests to see if payroll adjustments, like signing free agent talent or tanking, have a positive relationship with wins, and the second model tests to see if increasing payroll in general has a positive effect on wins. The coefficients for *delta\_payroll, payroll, delta\_attendance,* and *attendance* should all probably be positive as it would be expected wins would increase if payroll and attendance are high and increase. The coefficient for *previous\_wins* should be negative, since the more games a team wins the previous season, the harder it will be to win more the next season. The coefficient for *delta\_avgPay* should also be negative because an increase in payroll below the league average would be seen as a decrease relative to all of the other teams.

After running the regressions, those predictions turn out to be consistent with the results. In the first model Table 2, column 2, all variables and the constant are statistically significant at the 5% level, and all but the change in the average league payroll are statistically significant at the 1% level. Both *delta\_payroll* and *payroll* (rows 1 and 2) have positive coefficients. This means that teams can expect an increase of .1282 wins per million dollar increase, and an increase .0617 wins per million of total payroll. *Previous\_wins* (row 5) has a negative effect, so teams with more wins are less likely to see an increase in wins. *Delta\_avgPay* (row 6) also has a negative effect, meaning that increasing league wide payroll has a negative effect for teams that do not keep pace. In the second model in Table 2, column 3, *delta\_payroll*, *payroll*, and *previous\_wins* all remain statistically significant at the 1% level, as well as the new variable *delta\_attendance*. While *attendance* and *delta\_avgPay* are not statistically significant, they both still have strong p values and are significant at the 10% level. All coefficients from the previous regression still have the same signs from the previous model. Both attendance coefficients (rows 3 and 4) in the second model have positive signs this means that teams can expect an increase in wins if attendance is high and if it is increasing. The standard errors for both models are fairly consistent with each other, however, the t-stats are a little bit smaller for some of the variables in model 1 because some of the coefficients are smaller than those in model 2. The R2 in model 1 is .3878 and in model 2 it is .4569. This means that both models explain a fair amount of the variability in the change in wins. The F-stats for both are rather large, 56.22 and 49.50, meaning that the joint distribution is very strong. Time trends were left out of the models as it was both statistically insignificant and significantly lowered the t-stats of many of the other variables. Additionally, using log models did not appear to have any significant effect.

There were a few limitations for the models. The first being the control for revenue. Data for team revenue was difficult to find for all 12 years, so that is why attendance was used instead. While attendance can give good insight to revenue it is not perfect. It probably helps account for merchandise revenue as teams with higher attendance are likely to sell more merchandise. However, it will not be able to account for big changes in revenue as a result of TV deals, sponsorships, etc. Attendance also does not take into account differences in ticket prices, as Cubs tickets are much higher than Rays tickets. The model also doesn’t take into account any player metrics or injuries. So, a team with a lot of injured or underperforming players might have a biased result. Playoffs were excluded from the data because it would skew the data. Playoff teams play more games which means they will have the opportunity to win more games and increase attendance numbers, as well as increase other revenue sources. Additionally, teams tend to act differently in playoff games as the stakes are much higher.

Based on the data there is reason to conclude that increasing payroll does have a positive effect on increasing wins between seasons. Ignoring changes in revenue, teams can expect to win an additional game per $8 million increase in payroll. Given how large the payroll increase needs to be, this strategy is not too effective. Teams looking to make major gains by signing lots of free agents should expect to see smaller increases in wins, rather than major spikes in wins. Additionally, teams trying tank should not expect to become the worst team overnight. Rapidly increasing payroll would be a much better strategy for an 85-win team trying to make a push for the playoffs the next year (88 wins being a common cut off) than for a 70-win team trying to make the playoffs.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table 1 Summary Statistics | | | | |
| Variable | Mean | Standard Deviation | Min | Max |
| delta\_wins | 0 | 11.34 | -33 | 31 |
| delta\_payroll | 5.18 | 18.58 | -78.46 | 121.16 |
| payroll | 97.31 | 42.17 | 15 | 253.64 |
| delta\_attendance | 17.03 | 299.93 | -1,175.12 | 1,943.57 |
| attendance | 2,493.92 | 672.16 | 748.55 | 4,298.66 |
| previous\_wins | 80.99 | 11.22 | 43 | 105 |
| delta\_avgPay | 5.19 | 2.96 | -1.04 | 9.32 |

|  |  |  |
| --- | --- | --- |
| Table 2 Regressions of Change in Team Wins (delta\_wins) | | |
| Independent Variables | Model 1 | Model 2 |
| delta\_payroll | 0.1282 | 0.0930 |
| (0.0272)\*\*\* | (0.0277)\*\*\* |
| payroll | 0.0617 | 0.0484 |
| (0.0137)\*\*\* | (0.0177)\*\*\* |
| delta\_attendance |  | 0.0090 |
|  | (0.0016)\*\*\* |
| attendance |  | 0.0020 |
|  | (0.0011)\* |
| previous\_wins | -0.6304 | -0.6806 |
| (0.0475)\*\*\* | (0.0466)\*\*\* |
| delta\_avgPay | -0.3527 | -0.3179 |
| (0.1666)\*\* | (0.1639)\* |
| Constant | 46.2105 | 46.5275 |
| (3.6066)\*\*\* | (3.5103)\*\*\* |
| Summary Statistics | | |
| R2 | 0.3878 | 0.4569 |
| Observations | 360 | 360 |
| F-stat | 56.22 | 49.50 |
| Notes \*Significant at the 10% level \*\*Significant at the 5% level \*\*\*Significant at the 1% level | | |

Bibliography

Cohen, Gary. "Baseball Attendance." The Baseball Cube. N.p., 15 Apr. 2017. Web.

15 Apr. 2017. <http://www.thebaseballcube.com/extras/attendance/>.

Cohen, Gary. "Baseball Payrolls History." The Baseball Cube. N.p., 15 Apr. 2017. Web.

15 Apr. 2017. <http://www.thebaseballcube.com/extras/payrolls/>.

Hassan, Shahriar. "Can Money Buy Success?: A Study of Team Payroll and ..." The Journal of

Global Business Management, Oct. 2008. Web. 21 Apr. 2017.

"Major League Baseball Team Win Totals." Baseball-Reference.com. Sports Reference LLC, 15

Apr. 2017. Web. 15 Apr. 2017.

Schwartz, Noah L., and Jason M. Zarrow. "An Analysis of the Impact of Team Payroll on

Regular..." Undergraduate Economic Review. Illinois Wesleyan University, 2009. Web. 21 Apr. 2017.