Economics and Business Letters

7(2), 62-69, 2018

Team payroll, pitcher and hitter payrolls and team performance: Evidence from the U.S. Major League Baseball

Xing Lu[1] • Jason Matthews[2] • Miao Wang[3] • Hong Zhuang[1*]

[1] *Judd Leighton School of Business and Economics, Indiana University South Bend, South Bend, USA*
[2] *Gäshawk, South Bend, USA*
[3] *Department of Economics, Marquette University, Milwaukee, USA*

Received: 5 February 2018
Revised: 21 February 2018
Accepted: 23 February 2018

Abstract

We use the U.S. Major League Baseball team level data in 1985-2015 and find an inverse U-shaped relationship between team payrolls and winning percentages. Furthermore, when investigating the winning effects of pitcher and hitter payrolls, we find the similar curvilinear relationship between pitcher/hitter salaries and team performance.

Keywords: Major League Baseball; payroll; performance
JEL Classification Codes: L20, L83

1. Introduction

As a multi-billion industry, the average team payroll of Major League Baseball (MLB) has increased quickly from $10.6 million in 1985 to $121.9 million in 2015 with a wide variation in payrolls across MLB teams. In 2015, Los Angeles Dodgers ranked at the top with a payroll of $272.8 million, while Miami Marlins' payroll was at the bottom with $68.5 million. An interesting question emerges to baseball team management and fans: does the team with a bigger payroll have a better on-field performance?

A number of studies in the baseball economics literature suggest that team payroll is an important factor forecasting team performance. Wealthy teams can attract the best players by offering the highest salaries in a free agency market and thus assemble a high quality team to achieve better performance (Hall et al., 2002).[1] Often, a team's winning percentage is modeled as a linear function of the team's payroll in prior studies (Jane, 2010; Annala and Winfree, 2011).

[*] Corresponding author. E-mail: zhuangh@iusb.edu.

Citation: Lu, X., Matthews, J., Wang, M., and Zhuang, H. (20XX) Team payroll, pitcher and hitter payrolls and team performance: Evidence from the U.S. Major League Baseball, *Economics and Business Letters*, 7(2), 62-69.

[1] Hall et al. (2002) find that MLB team payroll had an insignificant effect on winning record in 1980-1992 but exhibited a significantly positive effect on performance in 1993-2000. Hall et al. (2002) also find that higher payroll is significantly associated with better performance using English soccer data in 1979-1999.

ISSN: 2254-4380

However, teams with largest payrolls do not always have the highest winning percentages. Unequal distribution of salaries resulted from some ‘star’ players receiving a significant portion of the payroll may discourage other players to provide effort and therefore undermines the overall team performance. Empirical studies by Wiseman and Chatterjee (2003), Jane (2010), and Annala and Winfree (2011) indicate a negative relationship between payroll inequality and winning percentages. Furthermore, similar to the idea of a backward-bending labor supply curve, initially players are willing to work harder for extra compensation at the low income level and thus the team performs better on field. However, when income continues to rise, the income effect might dominate the substitution effect that players would choose more leisure and thus, less effort, thereby weakening team performance (Scott et al., 2005). In this sense, the payroll–performance relationship may not be strictly linear even after controlling for within-team payroll disparity.

Our paper studies the relationship between MLB team payrolls and team performance and it contributes to the literature in two ways. First, we re-examine the baseball payroll-performance relationship in a non-linear setting and also take into account that team payrolls are potentially endogenous, both of which have not been addressed in prior studies. We attempt to shed light on baseball club management by providing better insights on the influence of team spending on winning percentages. Second, we estimate the winning effects of pitcher and hitter salaries to explore the possible differential effects of defensive and offensive capacities on team performance. Our paper is among the first ones to provide evidence on differential effects of payment to defensive and offensive capacities on team performance that contributes to team management.

Using data on 31 MLB teams in the U.S. over 1985-2015, we find an inverse U-shaped relationship between team payrolls and team performance after controlling for endogeneity as well as team and time fixed effects. A similar hump relationship between payroll and team performance is also observed among pitchers and hitters.

2. Methods

We estimate the effect of payroll of team *i* in year *t* on MLB team performance as follows:

$$Win\_percentage_{it} = \beta_0 + \beta_1 Payroll_{it} + \beta_2 Payroll^2_{it} + \beta_3 Gini_{it} + \delta_i + \gamma_t + \varepsilon_{it} \quad (1)$$

where $Win\_percentage_{it}$ is the percentage of winning games over total games played in a season; $Payroll_{it}$ is the sum of all players’ salaries in millions of dollars deflated by consumer price index (CPI); $Payroll^2_{it}$, the squared team payroll, is included to explore the possible curvilinear relationship between a team’s payroll and its performance; $Gini_{it}$ is the Gini coefficient measuring the disparity of team salary distribution. We also control for team specific fixed effects, δ_i, and year fixed effects, γ_t. Lastly, ε_{it} is the random error.

The variable *Payroll* is potentially endogenous as there might be two-way causality between team payrolls and team performance. Higher payrolls can lead to a team of better talents which increases the potential to win. When teams win more, they tend to generate more revenue (Scully, 1974; Bradbury, 2010). As club owners have access to more financial capital, they might be willing to pay more to attract talented players or retain current outstanding free agents in order to maintain their competitiveness. Furthermore, the findings in Hall et al. (2002) that team performance Granger causes wages using the MLB data over 1980 – 2000 also suggest the possible causality from performance to payroll. To address the potential endogeneity concern, we follow Jewell and Molina (2004) and include mean team age, the squared mean team age, market size proxied by the metropolitan population where the team is located, and a dummy variable for whether the team plays in the National League as instruments for the team payroll. The use of mean team age and its squared term is consistent with the Mincer-type earnings relationship. A large market size may help a team to generate more revenue and pay higher salaries but may not directly affect its winning percentage. The dummy variable controls for the possible difference in team salary structure between the National League and the American League.

To explore the offensive and defensive contributions to team victories, we define pitchers as defensive players and hitters as offensive players, following Scully (1974). While this method of categorizing is not perfect because there is no clear definition of defensive or offensive position in MLB, it is an efficient way to capture most players that fall into these two categories. Then we aggregate salaries of pitchers and hitters by team to generate variables *Pitcher_payroll* and *Hitter_payroll*, which are also converted to real values using CPI. We will substitute *Payroll* in Eq. 1 by *Pitcher_payroll* and *Hitter_payroll* respectively as well as their squared term and present the estimated results in Section 4.

3. Data

All baseball related data are obtained from the website of the Lahman Baseball Database2. The winning percentage (*Win_percentage*) is constructed by taking the wins of each team divided by total games played in a year multiplied by 100. The winning percentage takes a value between zero and 100. Team payroll, pitcher payroll or hitter payroll is an aggregate of salaries over all players, pitchers or hitters measured in millions of dollars. The salary data are open day salaries. In the case that players may be hurt, traded, or called up during a season and thus are not able to play, the salaries of players who do not play a single game in a season are not included in the total team salary. Furthermore, to rid the impact of inflation, all payroll variables are divided by CPI and converted to their real value. The CPI data are obtained from the Bureau of Labor Statistics, with a base year of 1982-1984. The Gini coefficient measuring the disparity of player salary distribution is calculated by authors for each team in every year. The Gini coefficient is an index between zero and one. Data on the U.S. metropolitan population are obtained from the Bureau of Economic Analysis. For Canadian teams, the data on metropolitan population are from Statistics Canada.

Table 1. MLB team list.

Arizona Diamondbacks	Minnesota Twins
Atlanta Braves	Montreal Expos
Baltimore Orioles	New York Yankees
Boston Red Sox	New York Mets
Chicago White Sox	Oakland Athletics
Chicago Cubs	Philadelphia Phillies
Cincinnati Reds	Pittsburgh Pirates
Cleveland Indians	San Diego Padres
Colorado Rockies	Seattle Mariners
Detroit Tigers	San Francisco Giants
Houston Astros	St. Louis Cardinals
Kansas City Royals	Tampa Bay Rays
Los Angeles Angels of Anaheim	Texas Rangers
Los Angeles Dodgers	Toronto Blue Jays
Miami Marlins	Washington Nationals
Milwaukee Brewers	

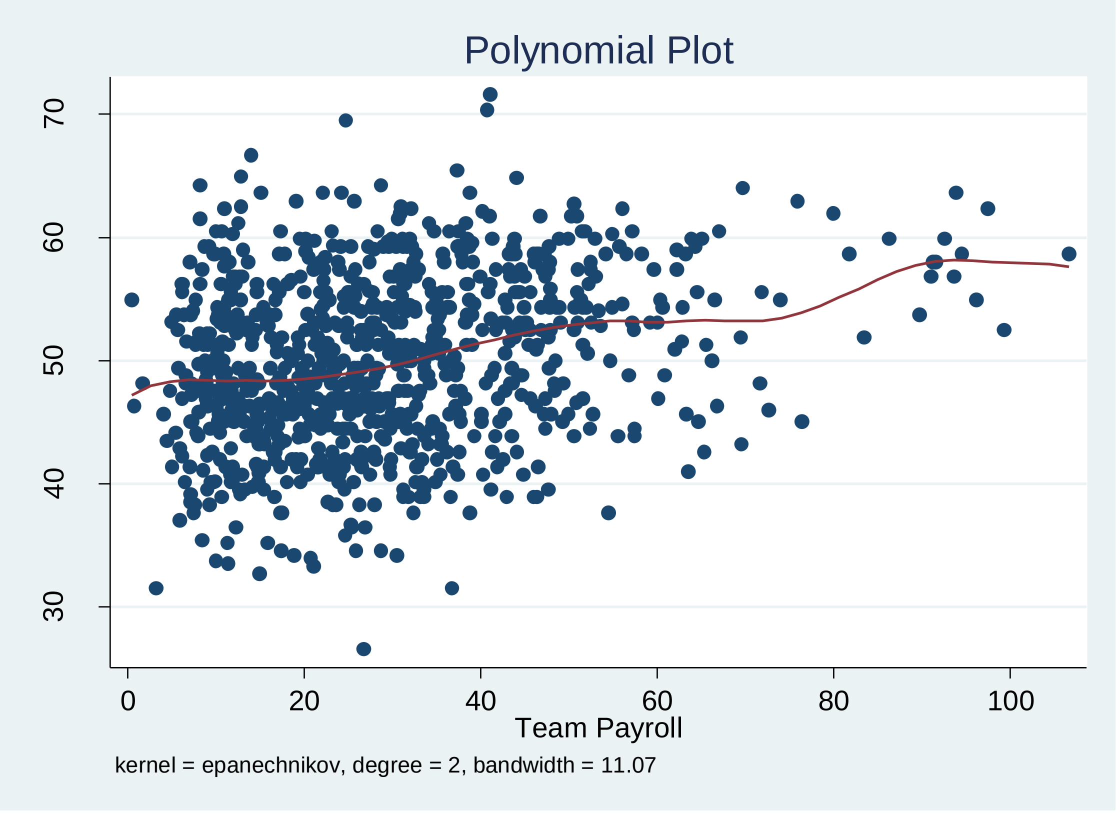
[2] http://www.seanlahman.com/baseball-archive/statistics/.

Table 2. Descriptive statistics.

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Win_percentage (%)	888	49.985	6.880	26.543	71.605
Payroll ($million)	888	29.167	17.273	0.450	106.676
Gini (units)	888	0.561	0.074	0.286	0.750
Pitcher_payroll ($million)	888	13.413	8.548	0.546	49.486
Hitter_payroll ($million)	888	17.514	10.922	0.229	66.960

Determined by data availability, our sample covers 31 MLB teams over 1985 - 2015. The sample is unbalanced. We provide the list of teams in Table 1 and descriptive statistics in Table 2. A scatter plot of polynomial estimates of winning percentage against team payroll is presented in Figure 1, which shows that winning percentage trends up as team payroll increases but starts to decline when team payroll is about $90 million.

Figure 1. Polynomial plot of winning percentage and team payroll.



4. Results

Fixed-effects estimated results of Eq. 1 are displayed in columns (1) and (2) in Table 3. Both regressions show that a team's payroll is an important factor to predict its winning percentage, which echoes findings in previous studies. Furthermore, our results show strong support for the non-linear relationship between team payroll and team performance. The coefficient on payroll is positive and significant at the 1% level and the coefficient on squared payroll is negative and significant at the 5% level or better in both regressions. Regression 3.2 suggests that for an average MLB team in 2015 with a payroll at $49 million, a $1 million rise in team payroll will increase the team winning percentage by 0.3 percentage point, which is approximately an increase of 0.5 game in wins, on average.[3,4] In other words, to win one additional game, the team needs to increase its payroll by $2 million, a 4.1% increase from the sample average MLB team payroll in 2015. These results provide a possible explanation why MLB teams continue to pay higher salaries to their players. Furthermore, regression 3.2 indicates that the positive effect of an increase in payroll on team performance peaks at about $87.6 million, consistent with our observation from Figure 1. Our results imply that initially, offering higher salaries to attract talents would have a positive impact on team performance. However, beyond the inflection point, there could be a drop in winning percentages associated with higher payrolls.

Table 3. Regression results.

	3.1	*3.2*	*3.3*	*3.4*	*3.5*	*3.6*	*3.7*	*3.8*
	FE	*FE- IV*	*FE*	*FE*	*FE*	*FE-IV*	*FE*	*FE-IV*
Payroll	0.425***	0.701***	0.352**	0.409***				
	[0.065]	[0.131]	[0.171]	[0.083]				
Payroll²	-0.003***	-0.004**	-0.002	-0.003***				
	[0.001]	[0.002]	[0.002]	[0.001]				
Top_Payroll			-0.043					
			[0.348]					
Gini	-20.096***	-21.407***	-12.629	-20.137***	-17.734***	-17.162***	-20.846***	-27.086***
	[4.811]	[5.027]	[10.043]	[6.262]	[5.044]	[5.173]	[5.145]	[5.419]
Pitcher_payroll					0.659***	1.414***		
					[0.116]	[0.332]		
Pitcher_payroll²					-0.009***	-0.015*		
					[0.002]	[0.009]		
Hitter_payroll							0.393***	1.282***
							[0.086]	[0.265]
Hitter_payroll²							-0.004***	-0.014**
							[0.001]	[0.006]
kleibergen-Paap underidentification test: p value	–	0.0000	–	–	–	0.0018	–	0.0000
Hansen overidenti-fication test: p value	–	0.1946	–	–	–	0.6414	–	0.2627
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observation	888	888	303	585	888	888	888	888
R-squared	0.1097	0.0454	0.1236	0.1352	0.0785	0.0994	0.0612	0.1115

Note: Robust standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1.

[3] The average number of games played in a season over the sample period is 160; as a result, a 0.3 percentage point increase in winning percentage is equivalent to winning an additional 0.48 game, which is obtained by 0.003*160=0.48.

[4] Test statistics suggest that the instruments are valid because the null hypothesis of the kleibergen-Paap under-identification test is rejected at the 1% level and that of the Hansen over-identification test cannot be rejected at conventional levels.

The coefficient on the Gini coefficient is negatively significant, suggesting that unequal distribution of salary is associated with a lower team performance. This is consistent with the argument that greater salary inequality can lead to lower team productivity and less cooperation among players and hence worsen team performance. One could argue that the negative effect of salary dispersion on winning percentage could be attributed to some high-paid players producing at a level that does not justify their pay, rather than less cooperation among team members. To further explore this, we consider a high-paid player as a player whose salary is three standard deviations above the team average. In regression 3.3, we create the variable *Top_Payroll* as the sum of high-paid players' salaries and *Payroll* is redefined as the sum of other players' salaries in the team.[5] Meanwhile, Payroll^2 is still the squared team payroll including all players. The fixed-effects estimation shows that *Top_Payroll* has an insignificant effect while the coefficient on payroll of the rest of players remains significantly positive. In the meantime, the coefficient of Gini coefficient becomes insignificant. For comparison, we re-estimate Eq. 1 using observations from teams that do not have high-paid players and report findings in regression 3.4. Results in regression 3.4 are similar to those in regressions 3.1 and 3.2, confirming the hypothesis that greater salary dispersion reduces within team cooperation and thereby negatively affects team winning percentage. Indeed, the findings in regression 3.3 also suggest that a significant share (about 21%) of team salary paid to the superstars may not always be worthwhile considering the insignificant effect of high-paid players' salary on wins.[6]

Results of the winning effects of defensive and offensive payrolls on team performance are reported in regressions 3.5 – 3.8.[7] Both positions contribute positively to team winning percentages and a similar hump shape relationship exists between each position's payroll and team performance. In particular, regressions 3.6 and 3.8 suggest that a $1 million increase in pitcher payroll will increase the winning percentage by 1.01 percentage points on average and a similar increase in hitter payroll is associated with a 0.79 percentage point improvement in winning percentage on average.[8] Further computation reveals that the inflection point associated with pitcher payroll is around $47.13 million and the inflection point associated with hitter payroll is about $45.79 million.

It is well known that baseball salaries have increased much faster than other prices. As a robustness check, we replace our payroll variable by a typified variable (Z_{Payroll}) in our regressions. The typified payroll is constructed by subtracting the average payroll of all teams in year *t* from the payroll of team *i* in year *t* and then dividing the difference by the standard deviation of payroll for all teams in year *t*. Similarly, we transform pitcher and hitter payrolls to their corresponding typified variables as well. We substitute *Payroll* and squared *Payroll* in Eq. 1 by the typified payroll variable and its squared term. The fixed-effects and IV fixed-effects results using the typified payroll are presented in regressions 4.1 and 4.2 of Table 4.[9] The findings are qualitatively similar to those in Table 3 – initially, a rise in team payrolls will lead to better performance; but once beyond the inflection point, further rise in team payrolls tend to decrease the winning percentage. This curvilinear relationship, by and large, holds for the typified payroll of pitcher and hitter as displayed in regressions 4.3, 4.5 and 4.6, although the coefficient of squared typified pitcher payroll is not estimated precisely in regression 4.4.

[5] *Top_Payroll* and *Payroll* are both deflated by CPI.

[6] The mean salary for high-paid players is $5.20 million, accounting for 21% of the team salary; while the average team salary including high-paid players is $24.73 million.

[7] Regressions 3.6 and 3.8 employ the same instruments as in regression 3.2. The instruments are valid as well.

[8] These estimates are based on the sample averages of hitter and pitcher payrolls.

[9] The mean team age and its squared term, metropolitan population and the dummy variable for national league are indicated as weak instruments for typified payroll variables. To address this concern, we include the first and second lags of typified payroll and its quadratic term as potential instruments. The specific instruments included in the regression are determined by the kleibergen-Paap under-identification and Hansen over-identification tests.

We also note that the R^2 values in Table 4 have improved compared to those in Table 3, yet they are still relatively low. This could be due to the long time span of our sample. For example, the R^2 based on the full sample in Annala and Winfree (2011) is 0.1882 using the U.S. MLB data over 1985-2004, comparable in value to the R^2 in our regressions.

Table 4. Robustness check.

	4.1 *FE*	*4.2* *FE-IV*	*4.3* *FE*	*4.4* *FE-IV*	*4.5* *FE*	*4.6* *FE-IV*
$Z_{Payroll}$	2.438***	4.759***				
	[0.293]	[0.567]				
$Z^2_{Payroll}$	-0.296*	-0.391*				
	[0.177]	[0.234]				
Gini	-18.765***	-21.554***	-17.293***	-20.977***	-20.583***	-30.377***
	[4.458]	[5.557]	[4.501]	[5.787]	[4.500]	[5.934]
$Z_{PitcherPayroll}$			3.187***	2.974***		
			[0.528]	[0.972]		
$Z^2_{PitcherPayroll}$			-1.311*	-1.847		
			[0.670]	[1.447]		
$Z_{HitterPayroll}$					1.562***	5.414***
					[0.280]	[0.659]
$Z^2_{HitterPayroll}$					-0.317*	-1.095***
					[0.176]	[0.385]
kleibergen-Paap underidentification test: p value	–	0.000	–	0.000	–	0.000
Hansen overidenti-fication test: p value	–	0.1452	–	0.3290	–	0.7640
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Team fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	888	798	888	826	888	826
R-squared	0.149	0.0523	0.104	0.0917	0.0890	0.149

Notes: Robust standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1.

5. Conclusion

This study assesses the relationship between MLB team payrolls and winning percentages. Our findings suggest that initially when a team pays more to players, it can enhance its winning percentage. Yet, such a positive effect disappears after a threshold value of team payroll is reached. Our results remain robust after we control for possible endogeneity of a team’s payroll as well as to an alternative measure of team payroll. Furthermore, we look at pitcher and hitter payrolls to disentangle the winning effects of defensive and offensive capacities. We find a similar inverse U-shaped relationship between salaries for pitchers and hitters and winning percentages.

References

Annala, C. N., and Winfree, J. (2011) Salary distribution and team performance in Major League Baseball, *Sport Management Review*, 14(2), 167-175.

Bradbury, J. C. (2010) *Hot stove economics: understanding baseball's second season*, New York, NY: Copernicus Books.

Hall, S., Szymanski, S., and Zimbalist, A. S. (2002) Testing causality between team performance and payroll: The cases of Major League Baseball and English Soccer, *Journal of Sports Economics*, 3(2), 149-168.

Jane, W. J. (2010) Raising salary or redistributing it: a panel analysis of Major League Baseball, *Economics Letters*, 107(2), 297-299.

Jewell, R. T., and Molina, D. J. (2004) Productive efficiency and salary distribution: the case of US Major League Baseball, *Scottish Journal of Political Economy*, 51(1), 127-142.

Scully, G. W. (1974) Pay and performance in Major League Baseball, *American Economic Review*, 64(6), 915-930.

Scott, J., Anderson, R., and Webb, J. (2005) The labor–leisure choice in executive compensation plans: does too much pay reduce REIT performance?, *Journal of Economics and Business*, 57(2), 151–163.

Wiseman, F., and Chatterjee, S. (2003) Team payroll and team performance in Major League Baseball: 1985-2000, *Economics Bulletin*, 1(2), 1-10.