



University of
Connecticut

Department of Economics Working Paper Series

The Interleague Advantage: A Difference in Differences Analysis

Brian Volz
University of Connecticut

Working Paper 2009-32

October 2009

341 Mansfield Road, Unit 1063
Storrs, CT 06269-1063
Phone: (860) 486-3022
Fax: (860) 486-4463
<http://www.econ.uconn.edu/>

This working paper is indexed on RePEc, <http://repec.org/>

Abstract

It has been argued that the introduction of interleague play in Major League Baseball provides an advantage to American League teams due to their use of the designated hitter. This paper examines whether this advantage actually exists and if so how large any advantage may be. The question is analyzed using a difference in differences model based on player performance data on interleague games from 1997 to 2008. It is shown that American League teams do have a small and statistically significant offensive advantage during interleague play. American League teams are estimated to have a 1.1 to 7.3 point advantage in batting average, a 0.1 to 9.8 point advantage in on base percentage, and a 1.2 to 9.9 point advantage in slugging percentage.

Journal of Economic Literature Classification: L83

Keywords: Baseball, League Structure, Difference in Differences

Introduction

Throughout the history of Major League Baseball teams have competed in two separate leagues. Traditionally, the winners of the American and National Leagues have met in the World Series each year to determine a champion. Due to this two league structure, prior to 1997, American League and National League teams did not compete against each other except for in this one best of seven championship series.

In 1997, Major League Baseball decided to schedule a limited number of interleague games during the regular season. From 1997 to 2001, teams competed only against their corresponding divisions. For example, the National League East would only play games against the American League East. In 2002, interleague play was expanded to include games between teams from various divisions while maintaining interesting matchups between local rivals. Despite complaints from baseball traditionalists, these interleague matchups have proved very popular with fans and therefore are expected to continue as a permanent part of Major League Baseball.

In addition to those who are in favor of maintaining baseball's traditions some fans are against interleague play because it is perceived to give an offensive advantage to American League teams. This hypothetical advantage comes from the fact that the American and National Leagues do not play by exactly the same rules. In the National League, pitchers are treated the same as all other positions and as such are expected to bat. However, in the American League a designated hitter bats in place of the pitcher. During interleague play the game is conducted according to the rules of the home team's league.

The presence of the designated hitter in the American League may give American League teams an advantage when competing against National League teams. This is due to the fact that National League teams are constructed based on having only eight batters starting each game. When National League teams play at an American League stadium they are forced to take a player who is normally a reserve and use them as a ninth batter in their starting lineup. American League teams on the other hand are constructed based on having nine starting batters. Therefore, when playing an interleague game by American League rules National League teams may face a disadvantage as they must compete with eight regular starters and one reserve against nine regular starters.

When playing interleague games at a National League stadium, American League teams are forced to have their pitchers bat. Due to their small number of plate appearances it is unlikely that American League pitchers will be as successful batting as National League pitchers. Therefore, American League teams may face some disadvantage when playing on the road. However, the offensive expectations for pitchers are very low even in the National League. Given the small contribution of pitchers to the offense any advantage enjoyed by National League teams is likely to be very small.

While American League teams may face a disadvantage from pitchers batting, that disadvantage may be offset by a superior bench. When American League teams play on the road they are forced to bench one of their usual starting players. Therefore, the American League team will have a quality starting player on their bench for use in strategic situations. It is likely that this player will be superior to any reserve player which the National League team may have. The difference in offensive skill between these two reserve players may be large enough to offset

the difference between pitchers. If this is the case the American League team will also enjoy an offensive advantage when playing at a National League stadium.

The goal of this paper is to determine if this offensive advantage actually exists and if so how large this advantage may be. Bradbury and Drinen (2006) exploit the existence of interleague play in order to evaluate the effect of the designated hitter on hit batsmen. However, they do not address the question of whether the presence of the designated hitter gives either team an advantage. Butler (2002) shows that interleague play results in a seven percent increase in attendance over a comparable intraleague game. However, Butler also does not address any possible advantage arising from the difference in rules. The question of whether interleague play provides either league with an offensive advantage appears to have been neglected in the literature prior to this analysis. The question is analyzed here using a difference in differences model based on player performance data on interleague games played from 1997 to 2008. The specifics of the methodology and data are presented in the following section. The results and their implications are then presented and discussed in the last two sections.

Methodology

If American League teams experience an offensive advantage during interleague play then this advantage should be apparent in the difference between the teams' lineups. However, having a superior starting lineup does not necessarily imply an advantage. This is due to the fact that the players not affected by the interleague rules may be of differing quality. If lineups can be observed under both sets of rules then any advantage which is independent of the rules can be differenced out. By eliminating these differences the overall advantage associated with interleague play can be identified.

More specifically, the question is modeled as follows. The difference in overall lineup quality is composed of two parts. The first part is the difference in quality attributable to those players unaffected by the rules. The second part is the difference in quality which is caused by the rules of that particular game. Therefore, the differences in offensive quality are modeled as follows. Equation (1) represents the model for games played under American League rules while equation (2) represents the model for games played under National League rules.

Game Played Under American League Rules:

$$(1) \text{Lineup}_{AL} - \text{Lineup}_{NL} = \text{Difference in Player Quality} + \text{Advantage}_{AL}$$

Game Played Under National League Rules:

$$(2) \text{Lineup}_{AL} - \text{Lineup}_{NL} = \text{Difference in Player Quality} + \text{Disadvantage}_{AL}$$

Lineup_{AL} and Lineup_{NL} are measures of the overall performance of the two team's lineups. The difference of these two measures is equal to Difference in Player Quality plus Advantage_{AL} or Disadvantage_{AL} . Difference in Player Quality is the difference in offensive performance independent of the rules. Advantage_{AL} and Disadvantage_{AL} are then the differences attributable to the rules of that particular game.

In order to identify the overall advantage or disadvantage of American League teams in interleague play the difference of the two right hand sides can be taken as follows.

Difference in Differences:

$$(3) (\text{Difference in Player Quality} + \text{Advantage}_{AL}) - (\text{Difference in Player Quality} + \text{Disadvantage}_{AL}) = \text{Advantage}_{AL} - \text{Disadvantage}_{AL}$$

The resulting $(\text{Advantage}_{AL} - \text{Disadvantage}_{AL})$ measures the difference between the advantage American League teams face at home and the disadvantage they face on the road. A positive value for $(\text{Advantage}_{AL} - \text{Disadvantage}_{AL})$ implies that the American League enjoys an

offensive advantage attributable to differences in league rules. This advantage may be the result of two different circumstances. The first circumstance is that the American League team may enjoy an advantage at home and away. Alternatively, the home advantage may simply outweigh any disadvantage on the road. In either case a positive value implies that the American League will have an overall offensive advantage in interleague competition.

A negative value for $(\text{Advantage}_{\text{AL}} - \text{Disadvantage}_{\text{AL}})$ implies that the National League enjoys an offensive advantage attributable to differences in league rules. This advantage may also be the result of two circumstances. Again, either the National League team enjoys an advantage at home and away or the home advantage outweighs any disadvantage on the road.

A third possibility is that $(\text{Advantage}_{\text{AL}} - \text{Disadvantage}_{\text{AL}})$ is equal to zero. This implies one of two possibilities. The first possibility is that the advantage at home is completely offset by a disadvantage on the road. Alternatively, a value of zero may be the result of neither team having an advantage in either location. The implications of these different results are summarized in Table 1.

Table 1

Interpretation of Results

Result	Implication	Implication in Words
$(\text{Advantage}_{\text{AL}} - \text{Disadvantage}_{\text{AL}}) > 0$	$\text{Advantage}_{\text{AL}} > \text{Disadvantage}_{\text{AL}}$	AL home advantage outweighs NL home advantage or AL has advantage under both rules.
	Or $\text{Advantage}_{\text{AL}} > 0 \ \& \ \text{Disadvantage}_{\text{AL}} < 0$	
$(\text{Advantage}_{\text{AL}} - \text{Disadvantage}_{\text{AL}}) = 0$	$\text{Advantage}_{\text{AL}} = \text{Disadvantage}_{\text{AL}}$	AL home advantage is equal to NL home advantage or neither league has an advantage
	Or $\text{Advantage}_{\text{AL}} = 0 = \text{Disadvantage}_{\text{AL}}$	
$(\text{Advantage}_{\text{AL}} - \text{Disadvantage}_{\text{AL}}) < 0$	$\text{Advantage}_{\text{AL}} < \text{Disadvantage}_{\text{AL}}$	NL home advantage outweighs AL home advantage or NL has advantage under both rules
	Or $\text{Advantage}_{\text{AL}} < 0 \ \& \ \text{Disadvantage}_{\text{AL}} > 0$	

In order to identify $(\text{Advantage}_{\text{AL}} - \text{Disadvantage}_{\text{AL}})$ data must be available for both equation (1) and equation (2). In order for Difference in Player Quality to cancel out in equation

(3) the teams being compared must be the same under both (1) and (2). In other words, to identify the rules advantage the same teams must be observed playing under both National and American League rules. Luckily, during interleague play the same teams often play a series of games at each stadium. Therefore, a pair of games can be identified and used to isolate the rules advantage.

For example, if the Philadelphia Phillies play the Boston Red Sox both at home and away these two games can be used to identify the rules advantage based on this difference in differences model. The difference between the Phillies and Red Sox lineups can be taken both in Philadelphia and in Boston. These two differences can then be subtracted to identify any advantage from interleague play.

This concept relies on the assumption that the players unaffected by the rules change do not differ in quality from one game to another. If this is the case then the differences in quality unrelated to the rules change will cancel out leaving only the change in quality attributable to the rules. It will be shown in the following section that this assumption is indeed reasonable for the data chosen.

Data

In interleague play it is common for opposing teams to play a series of games at each stadium. Usually these series consist of two or three games at the National League stadium and two or three games at the American League stadium. These series provide an excellent opportunity to identify any advantage due to interleague play.

For this analysis all interleague series which consist of games at both stadiums within the same season are identified. From these series each National League game is paired with each

corresponding American League game. For example, a two game series played at each stadium between the Phillies and Red Sox results in four pairs of games. This is because the first game in Philadelphia is paired with both games in Boston and then the second game in Philadelphia is also paired with both games in Boston. Similarly, a three game series at each stadium will result in nine pairs of games. These pairs are constructed for every interleague series played at both stadiums from 1997 to 2008. This results in a sample of 1,117 pairs of games.

In order to identify any offensive advantage due to differing rules the quality of the starting lineups for each team must be compared for each pair of games. The starting lineups for each team are identified through game logs provided by Retrosheet¹. In order to quantify the quality of each lineup offensive statistics from that year are computed for each player on each team. Unfortunately, there is no one statistic which captures the overall offensive performance of a player. Therefore, the analysis is conducted using several commonly used offensive performance measures. These include batting average (AVG), on base percentage (OBP), slugging percentage (SLG), and on base plus slugging (OPS).

Pitchers, unlike batters, do not play every game. Therefore, their offensive statistics are based on a much smaller sample size. This is especially true for American League pitchers who only bat during interleague games played at the National League stadium. This can result in a pitchers offensive statistics being based on only a handful of observations. Such a small sample can result in statistics which do not represent the offensive skill of that player. In order to avoid this problem pitchers offensive quality is not based on their statistics from that season. Rather, their career offensive performance measures are included in the analysis.

The average offensive quality of each lineup is calculated based on that season's offensive statistics for each position player and career statistics for each pitcher. This measure is calculated for AVG, OBP, SLG, and OPS. The difference in differences is then taken for each pair of teams under each performance measure. The results of these calculations are presented in Table 2 and are discussed in the next section.

Table 2

Difference in Difference Results: Full Sample 1997-2008

Offensive Measure	Mean		Standard Error	95% Confidence Interval	
Batting Average	0.00280	***	0.00073	0.00137	0.00422
On Base Percentage	0.00409	***	0.00105	0.00203	0.00614
Slugging Percentage	0.00394	***	0.00111	0.00176	0.00612
On Base Plus Slugging	0.00770	***	0.00174	0.00429	0.01112
n	1,117				

***Significantly different from zero at the 1% level

**Significantly different from zero at the 5% level

*Significantly different from zero at the 10% level

In order for these difference in differences to capture the interleague advantage it must be the case that the quality of the players not affected by the differing rules are similar in the games being compared. In order to test this hypothesis, the number of players which appear in both lineups of each pair is calculated. If the lineups are exactly identical except for the interchanging of a pitcher for a designated hitter then eight players should be the same on each lineup. The average number of identical players is over six for both American and National League teams. This provides strong evidence that the lineups are indeed similar between games played at each stadium. However, in order to be certain that the assumption of similar lineups between locations is satisfied difference in differences are also calculated based on a sample which is restricted to those games for which each team has at least seven identical players in each game.

The results for this restricted sample are provided in Table 3 and discussed in the following section.

Table 3

Difference in Difference Results: Restricted Sample 1997-2008

Offensive Measure	Mean		Standard Error	95% Confidence Interval	
Batting Average	0.00422	***	0.00157	0.00113	0.00731
On Base Percentage	0.00524	**	0.00232	0.00066	0.00982
Slugging Percentage	0.00557	***	0.00220	0.00123	0.00990
On Base Plus Slugging	0.00928	***	0.00370	0.00197	0.01659
n	166				

***Significantly different from zero at the 1% level

**Significantly different from zero at the 5% level

*Significantly different from zero at the 10% level

Results

As can be seen in Table 2 and Table 3, the mean difference in difference is positive and significantly different from zero for all performance measures using both the full and restricted samples. This implies that American League teams enjoy an offensive advantage during interleague play. It should be noted that these results can be interpreted as a lower bound on the American League advantage. These results are based on starting lineups and, therefore, do not capture any additional advantage which results from American League teams having a designated hitter on the bench in National League stadiums.

Based on the sample with only high quality matches, American League team's starting lineups will have a batting average advantage of 4.22 points during interleague play. In 2008, the average Major League team had 34.3 at bats per game. Therefore, over a six game interleague series an American League's starting lineup is expected to have an additional 0.87

hits due to their rules advantage. In 2008, each team played 18 interleague games. Therefore, American League teams are expected to have an advantage of 2.61 hits each season during interleague play. While this is not a large advantage 2.61 hits could be the difference in a game. Additionally, based on the 95 percent confidence interval, the advantage may be as high as 10.25 hits per season.

When considering on base percentage under the restricted sample, American League teams have an advantage of 5.24 points. Based on 2008 season averages, this implies that American League teams have a 1.2 base runner advantage over a six game interleague series. This translates to a 3.6 base runner advantage over all 18 interleague games. Based on the 95 percent confidence interval this advantage could be as high as 6.8 base runners per interleague season.

Based on the restricted sample, American League teams also have an advantage in slugging percentage of 5.57 points. Based on 2008 averages, this implies an advantage of 1.14 total bases per six game interleague series. This results in an advantage of 3.44 total bases over the course of interleague play each season. As with batting average and on base percentage, the total bases advantage may be much larger. Based on the 95 percent confidence interval American League teams may have an advantage of as many as 6.12 total bases per interleague season.

American League teams also have an advantage in on base plus slugging of 9.28 points under the restricted sample. On base plus slugging does not have a direct interpretation in terms of hits, base runners, or total bases per game. However, it is equal to the sum of on base percentage and slugging percentage for which results have also been presented. The average

team on base plus slugging for 2008 was .749. Based on this average American League teams will have a 1.2 percent higher on base plus slugging due to their interleague advantage. This advantage may be as large as 2.2 percent based on the 95 percent confidence interval.

Conclusions

Based on difference in differences taken for 1,117 pairs of interleague baseball games between 1997 and 2008 it is concluded that American League teams enjoy a small but significant offensive advantage in the quality of their starting lineup due to interleague play. These advantages are present for all four offensive statistics considered and are larger when the model is restricted to only the highest quality matches.

The American League advantages consist of a 4.2 point increase in batting average, a 5.2 point increase in on base percentage, a 5.6 point increase in slugging percentage, and a 9.3 point increase in on base plus slugging. All of these results are significantly different from zero at conventional levels. Additionally, these results may be viewed as a lower bound on the overall interleague advantage as they do not capture any advantages the American League may have from the presence of a starting player on their bench when playing by National League rules.

While these advantages appear small baseball has long been known as a game of inches and it is likely that these small advantages over time could lead to a disproportionate number of wins for American League teams. This idea appears to be supported by the fact that the American League has won 52 percent of interleague games from 1997 to 2008.

References

- Butler, M. R. (2002). Interleague Play and Baseball Attendance. *Journal of Sports Economics*, 3(4), 320-334.
- Bradbury, J. C. & Drinen, D. (2006). Research Notes: The Designated Hitter, Moral Hazard, and Hit Batters: New Evidence From Game-Level Data. *Journal of Sports Economics*, 7(3), 319-329.
- Interleague History. (2009). Retrieved June 6, 2009, from <http://mlb.mlb.com/mlb/history/interleague/index.jsp>
- Retrosheet Game Logs. (2009). Retrieved March 30, 2009, from <http://www.retrosheet.org/gamelogs/index.html>

Footnotes

¹The information used here was obtained free of charge from and is copyrighted by Retrosheet. Interested parties may contact Retrosheet at “www.retrosheet.org”.