Identifying Moral Hazard: A Natural Experiment in Major League Baseball

John Charles Bradbury*   Douglas Drinen**

Sewanee: The University of the South

July 2004

Abstract

In baseball, allowing a designated hitter (DH) to bat for the pitcher creates the potential for moral hazard among pitchers, who may then hit more batters without the fear of retaliation by the opposing team. The use of the DH in only one of Major League Baseball's two leagues provides a natural experiment to test for the existence moral hazard in a controlled setting. We develop a new micro-level dataset of individual plate appearances, which allows us to control for detailed cost-benefit attributes that affect the decision calculus of the pitcher. We find that moral hazard explains 60 to 80 percent of the difference in hit batsmen between leagues and find evidence of direct retaliation against plunking pitchers. (*JEL* D81, J28)

Thanks to David Smith and the volunteers of the Retrosheet organization for compiling and sharing the play-by-play data. We also thank Kevin Holman, Jim Porter, participants at the annual meeting of the American Mathematical Society and the Mathematical Association of America, and Baseball Primer readers for helpful comments and suggestions.

Contact Information:
*Department of Economics, The University of the South, Sewanee, TN, 37383. Phone: (931) 598-1721. E-mail: jbradbur@sewanee.edu.

**Department of Mathematics and Computer Science, The University of the South, Sewanee, TN, 37383. Phone: (931) 598-3370. E-mail: ddrinen@sewanee.edu.

"Are you seriously going to throw at somebody when you're facing Randy Johnson?"

-- Curt Schilling (Rosenthal, 2002)

I. Introduction

In 1973, amid waning fan interest in baseball, the American League (AL) of Major League Baseball (MLB) instituted the designated hitter rule as an experiment. Inadvertently, the architects of this baseball rule change created an excellent natural experiment of great interest to economists studying moral hazard. The stated goal of this rule change was to boost offensive output by increasing the talent pool of batters in the lineup. Traditionally, the competing teams field nine players who must play defense in the field and bat. Because of the importance of the pitcher, who is responsible for putting every ball in play, teams rely on pitching ability and ignore the hitting ability when choosing the pitcher in the lineup. Therefore, pitchers tend to be very poor hitters. By allowing teams to substitute a player of greater hitting ability (known at the designated hitter or DH) to bat for the pitcher the total offensive output should increase. The experiment has since grown into an institution in the AL that differentiates it from the National League (NL) where all players must bat and play in the field.

The DH succeeded in turning the AL into the “power league” as intended, but an unintended consequence of the rule change is that the AL now has more batters hit by pitches. Though it has long been part of traditional baseball lore, Goff, Shughart, and Tollison (1997, hereafter, GST) first formally hypothesized that a cost differential to pitchers for hitting batters, which created the opportunity for moral hazard, is the cause of the interleague difference in hit batsmen. Veteran NL manager Dusty Baker describes the phenomenon as understood by the players, "You can be bold in (the American) League and get away with [hitting batters]. It's different in our league where you have to hit” (Morgan, 2002). Pitchers who do not have to bat (where they might face retaliation) are more willing to risk hitting batters than pitchers who do bat. In economic terms, AL pitchers are actors who are improperly "monitored" by the opposition and therefore are more likely to engage the undesirable behavior of hitting batters. The teammates and manger of the pitcher, in addition to the opposing team, may suffer as a consequence of their shirking agent, creating the potential for moral hazard.[1] Given that the rules of the game in both leagues are virtually identical except for the use of the DH, MLB created ideal conditions to test for the existence of moral hazard in a controlled setting.[2] In the initial study of this subject, using annual league time series data from 1901-1990, GST analyzes changes in the difference in hit batsmen between leagues before and after the implementation of the DH. The results indicate that, controlling for several factors, the introduction of the DH raised the level of hit batsmen in the AL 15 percent higher than the NL, consistent with the moral hazard hypothesis.

However, though the evidence is clear that the DH is associated with an increase in hit batsmen, it is not necessarily a consequence of moral hazard. Trandel, White, and Klein (1998, hereafter, TWK) and Levitt (1998) are skeptical of the interpretation of GST on two grounds. First, adding the DH raises the marginal benefit of hitting an additional good batter in the line-up. Because pitchers are poor hitters, few teams want to risk hitting the other pitcher. Replacing the pitcher with the DH places a batter more worthy of plunking in the lineup. As TWK state, "the difference [in hit batsmen] is largely

[1] Teammates may experience retaliation and perform sub-optimally at the plate in fear of probability retaliation. The manager receives sub-optimal performance of the pitcher and players and is subject to fines, suspensions, and ejections. In this sense, the pitcher is an agent of the team who shirks his responsibility to his principals in response to imperfect monitoring.

[2] The controlled environment of sports has become a popular arena for testing economic hypotheses. Some examples include Duggan and Levitt (2002) on identifying corruption in sumo wrestling; Goff, McCormick and Tollison (2002) on racial integration and innovation as measured by winning in baseball and basketball; McCormick and Tollison (1984) and Heckelman and Yates (2003) on crime and enforcement in basketball and hockey; and Sobel and Nesbitt (2004) on offsetting behavior in response to safety regulations in NASCAR racing.

because the AL batters are (on average) better hitters, and are thus less costly and more beneficial to hit." Therefore, the aggregate increase in hit batsmen in the AL over the NL may be attributed to the change in the composition of the lineup and not a moral hazard response of AL pitchers.[3] Second, the moral hazard story seems implausible, because pitchers are rarely hit and therefore rarely bear any retaliatory costs from hitting batters. Hitting the pitcher in retaliation may actually reward the instigator by putting the weak-hitting pitcher on base. Given that pitchers know this, it is unlikely that pitchers consider reciprocal retaliation to be a cost. As evidence, Levitt (1998) finds that there is no correlation between pitchers hitting batters and pitchers being hit themselves, and pitchers are hit so rarely that they would have to be hypersensitive to the relative price change induced by the DH. Both of these studies conclude that the increase in the quantity of hit batsmen in the AL likely reflects compositional changes in the batting order, and not the moral hazard of pitchers.[4]

The crux of the problem lies with aggregate data with which researchers can observe only the increase in the quantity of hit batsmen with the DH, a result that both competing hypotheses predict. Aggregate data, though informative, only reveal a quantity hit without regard to different situations where the costs and benefits vary. A retaliatory plunking war in one game may be hidden within yearly statistics from a 162 game season, even though such an event may signal a credible threat of retaliation that persists into the future. Therefore, the differing hypothesized effects are hard to disentangle with aggregate data. Thanks to a new dataset, we are able to break out individual events from the noise of aggregate statistics. Thus, we examine the costs and

[3] In reply Goff, Shughart, and Tollison (1998) responds to the criticisms by referring to the control variables (batting and pitching statistics) in previous and revised regressions that control for compositional factors.

[4] These studies also extend the data into the 1990s, in which the difference in hit batsmen between leagues becomes less stable. However, there are several plausible explanations for the change that we discuss in Section III.

benefits of hitting any particular batter at the time the pitcher chooses a pitch. To do this we extracted plate appearance observations from publicly available play-by-play data over several years. This data allows us to examine pitcher decisions as costs and benefits change with individual plate appearances. Controlling for several additional factors not included in past studies, we find that, though compositional factors are important, pitchers are more likely to hit batters when the DH is in effect than when it is not. A second important finding is that, though pitchers are hit more rarely than other players, pitchers do experience retaliation for hitting batters. Therefore, we find the previously unidentified retaliation against pitchers, which is needed to generate moral hazard behavior by pitchers in the NL. These findings indicate that the moral hazard of pitchers induced by the DH does contribute to hit batters.

II. The Costs and Benefits of Hitting Batters

TWK highlights the importance of the team's cost-benefit calculation when a pitcher hits a batter. Though it is the pitcher who actually throws the ball, he is an agent of the manager and not a completely autonomous actor. Assuming the manager wants to maximize the chance of winning and the pitcher is an agent of the manager, instructing a pitcher to hit the opposing pitcher (typically the weakest batter on the team) is rarely a good short-run strategy. Instead, many managers may choose to retaliate against the pitcher's better hitting teammates. This cost-benefit calculation, which is known by both teams, means that the pitcher should have little fear of retaliation. Thus, the observed interleague difference in hit batsmen is likely due to in inclusion of better hitters in the lineup rather than the moral hazard of pitchers. We agree with TWK that the cost-benefit calculation is important; however, observing these calculations is not easy using summary statistics. Many factors other than moral hazard affect pitcher decisions for each plate appearance. Isolating these specific factors is important for our analysis so that we can properly identify the cause of the increase in hit batsmen in the AL following the introduction of the DH.

Getting hit by a pitch is a painful, and sometimes dangerous event in baseball.[5] To deter the pitcher from accidentally or purposely hitting batters the hit batter is awarded first base. Given this statutory constraint, we can observe the benefits and costs of hitting a batter. The benefits include: preventing a batter from hitting the ball, which might generate more runs than simple base advancement; inflicting injury to a competitor; or decreasing the opposing team's willingness to stand close to the plate to hit outside pitches. The costs include: putting a runner on base, which increases the chance of the opposing team scoring; ejections, fines, and suspensions to the pitcher or manager; physical retaliation via assault; retribution against the teammates of the pitcher by plunking them when they bat; and direct reciprocal retribution in the form of the current pitcher being intentionally plunked by the opposing pitcher.

It is this last cost which is the main subject of discussion. AL and NL pitchers bear identical costs and benefits when facing batters except for the reciprocal retribution. AL pitchers do not bat and therefore bear less of the cost of hitting a batter. GST hypothesizes this relative price difference induces pitchers to engage in moral hazard.[6] This is not to say that any pitcher actively seeks to consume hit batsmen purposely (at least not most of the time). Instead pitchers engage in riskier behavior (e.g., throwing inside, faster) when DHs bat in their place, which results in a greater number of hit

[5] Pitches hitting batters have ended several professional baseball careers; seven players (only one in MLB) have died from their injuries. For a sample of serious injuries resulting from players hit by pitches see James (1985).

[6] Two other unpublished studies of the relationship between the DH and hit batsmen also find support for the moral hazard hypothesis. Using differences in hit batsmen between relievers (who rarely bat in the NL) and starting pitchers (who do bat) who switch leagues, Rupert and Stephenson (2002) find some support for the moral hazard explanation. Kawaura and La Croix (2002) replicate the GST and TWK estimations on Japanese professional baseball, which also has the DH in only one league. The results are consistent with Goff, Shughart, and Tollison (1998) in support of the moral hazard hypothesis.

batsmen. Cy Young award winning pitcher Randy Johnson, who has played in both leagues, explains, "If you're the pitcher and you're playing in the American League, then you may have a tendency to throw inside a little bit more knowing when that ninth hole comes up, you won't be hitting. You're protected in that regard" (Rosenthal, 2002). Consistent with the law of demand, lowering the negative consequences of an action leads people to take more risks that result in the negative outcome. This is a behavioral response similar to automobile drivers in Peltzman (1975), which documents seat belt laws inducing motorists to engage in riskier behavior that leads to car accidents.

However, TWK and Levitt (1998) propose that this cost differential between leagues is too minor to have much, if any, effect on an NL pitcher's decision when he pitches to batters. *On average*, hitting the pitcher is not a good idea. Levitt (1998) provides data for the 1990s indicating that pitchers who have a propensity for hitting batters are not more likely to be hit themselves when compared to pitchers who rarely hit batters.[7] However, circumstances arise where the benefits of hitting a pitcher exceed the costs. This may occur in the course of the same game in which a pitcher hits a batter or it may occur in some future game.[8] Though pitchers are hit less than other position players, it is possible that pitchers may be hit more than they ought to be, given their hitting ability. Ultimately, this is an empirical question. Using the data we describe in the following section, we can answer the following questions central to the analysis. Do pitchers hit more batters with a DH rule in effect when controlling for composition

---

[7] Trandel (2004) finds similar results for the 1974-1977 seasons using the same test. In addition, the author finds little evidence of retaliation against any batter. Using team data from 1960-2002, there is no relationship between the rate at which a team hits batters and the rate at which the same team's batters are hit.

[8] A well-documented case of how long teams will wait to retaliate against a pitcher occurred in 2002. Shawn Estes of the New York Mets threw at (and missed) Roger Clemens in retaliation for hitting Mets' star Mike Piazza two seasons prior. Mets' manager Bobby Valentine clearly ordered the retaliation because Estes was not a Met at the time Clemens struck Piazza.

changes in lineups? Is a pitcher more likely to be plunked when he bats after hitting an opposing player?

III. The History of Hit Batters and the Designated Hitter

Figure 1 shows the fluctuation in the rate of hit batsmen from 1921-2003. It is evident from the figure that the introduction of the DH in 1973 coincides with the AL's continued higher rate of hit batsmen as compared to the NL. From the introduction of the DH until the present, the AL's average yearly hit batsmen rate has exceeded the NL rate by 15 percent. From the beginning of the modern era of baseball until the 1970s, the hit batsmen rates varied closely, yet randomly, across leagues. For the next 21 years the AL rate exceeded the NL rate every year, which is consistent with the moral hazard and batter composition hypotheses. But in the mid-1990s, just as economists began to study this issue, a deviation in this pattern emerged. In 1994, the NL hit batsmen rate rose above the AL rate for the fist time since the implementation of the DH, and the leagues' hit batter rates began to fluctuate more closely than in the past. Why was there such a dramatic shift and how does this shift impact the perceived relationship between the DH and hit batsmen? There are two important factors that contributed to this change: league expansion and a rule change in the punishment for pitchers hitting batters.

In response to criticism over the recent events of the 1990s, Goff, Shughart, and Tollison (1998) postulates that the NL expansion in 1993 diluted the talent pool of players increasing the number of accidental hit batsmen. We believe the evidence supports this contention. Fringe pitchers and hitters are more apt to make mistakes by hitting batters or getting hit by pitches.[9] In 1993, the NL hit batsmen rate fell to only 10 percent of the AL rate, while the average difference over the preceding two decades was 20 percent. Additionally, the mid-1990s corresponded with a dramatic increase in hit batsmen in both leagues, doubling over the decade. The expansion hypothesis is consistent with both the change in the difference between leagues and the rise in hit batters overall. In the two rounds of expansion, three of the four expansion teams joined the NL with one existing AL team (Milwaukee) shifting to the NL. But more importantly, the asymmetric expansion draft rules in 1993 favored expansion players coming from NL rosters. Therefore, the expansion diluted the talent pool of both leagues, but affected NL rosters to a greater degree than AL rosters.[10] From 1993 to 1997, as the expansion players eased into the league, the NL rate exceeded the AL rate in three of the five years, with the AL rate being 2 percent lower than the NL rate, on average. But in 1998, when both leagues added a team, there was no asymmetry between leagues in drafting expansion players. Interestingly, the disruption to the previous pattern was noticeably smaller. Since the last round of expansion in 1998, only once (2000) has the NL rate exceeded the AL rate, with the AL rate exceeding the NL rate by an average of 6.5 percent.

The second factor was a rule change in the punishment for hitting batters that significantly altered the cost of retaliating against pitchers who hit batters. In 1994, MLB adopted the "double-warning" rule, which authorizes umpires to warn *both* teams if an

[9] See Gould (1996) for an interesting theory of how the dispersion of talent causes the best players to excel and the worst players to suffer. Schmidt and Berri (2003) find some support for Gould's hypothesis in regards to the influx of new talent into MLB.

[10] In the 1993 expansion, AL teams were allowed to protect more players on their rosters than NL teams. See Pappas (1997) for a description of expansion draft rules.

umpire judges that a pitcher intentionally hits or attempts to hit a batter.[11] Once the umpire issues a warning, retaliation results in the immediate expulsion of the current pitcher and manager of the offending team, accompanied by monetary fines. This significantly raises the cost of retaliation, thereby lowering the cost to pitchers for hitting batters. With the double-warning rule, pitchers can take risks that are more likely to result in hit batters than without the rule, because pitchers know the opposition will be less likely to retaliate due to the increased penalty. The rise in the NL rate of hit batsmen is consistent with the moral hazard hypothesis, and therefore is expected. NL pitchers are now more protected from retaliation than before the double-warning rule, thereby inducing NL pitchers to behave more like their counterparts in the AL. A similar phenomenon occurred in Japan. Japanese professional baseball, which also has the DH in only one league, experienced a similar convergence of hit batsmen after several years of excessive hit batsmen in the DH league. Kawaura and La Croix (2002) find this convergence correlates with a rule change that increased the penalty for intentionally hitting batters.

Finally, using game-level (not play-by-play) data from 1973-2003, Bradbury and Drinen (2004) find games in which the DH is allowed are associated with more hit batters than non-DH games, controlling for many relevant factors. Over the entire 31-year history of the DH, the four years in which the NL rate exceeded the AL rate appear to be exceptions rather than the rule. When relevant historical facts are included in the analysis it is clear that the 1990s actually lend support to, rather than cast doubt upon, the moral hazard explanation for differences in hit batsmen between leagues. The most recent history of hit batsmen rates is indicative of a return to the pre-1990s league differences in hit batsmen, with a tempering of the difference by the double-warning rule.

---

[11] In December 1993, MLB adopted the *On-Field Behavior Policy*, which, along with several new policies, created the double-warning rule (*N.A.P.B.L. Umpire Manual*, 1996). We would like to thank Jim Porter for identifying the exact timing of this rule change.

IV. Empirical Model

A data gathering project known as Retrosheet has carefully reconstructed baseball play-by-play data for many seasons.[12] Using a computer program, we extracted data by plate appearance from the Retrosheet data. From this we can observe the exact game situation when a player is hit or not hit; thereby, we can control for the costs and benefits of hitting any particular player during a plate appearance.[13] We use a predictive model to estimate the likelihood that a batter is hit given five categories of influences on hitting batters: moral hazard (DH), batter quality (BQ), pitcher quality (PQ), retaliation(R), and game situation (GS). Thus,

$$(1)\quad HBP_j = \alpha_j + \beta DH_j + \lambda \mathbf{BQ}_j + \gamma \mathbf{PQ}_j + \phi \mathbf{R}_j + \psi \mathbf{GS}_j + \eta \mathbf{YR}_j + \varepsilon_j \,;$$

where HBP is a dummy variable equal to one when a pitcher hits a batter and zero otherwise. Subscript j represents a plate appearance, α is the constant, and ε is the error term. We estimate Equation 1 using logit and probit estimation techniques.

The variable DH is a dummy equal to one when the DH rule is in effect. According to GST, it is the moral hazard of the DH that causes hit batsmen differential between leagues. If, after accounting for the many factors we include, the DH coefficient remains positive, it is likely that moral hazard explains the difference. If it is insignificant when controlling for other factors, some predicted by TWK and Levitt and others we add, it is likely that the moral hazard hypothesis is incorrect.

---

[12] Turocy (2003) uses Retrosheet play-by-play data to test game-theoretic predictions regarding stolen bases.

[13] We use baseball designation "plate appearance" as opposed to "at bats," because several results of plate appearances (including being hit by a pitch) do not count as at bats.

The **BQ** vector includes two batter quality variables that affect the cost-benefit calculation of hitting batters. Better hitting batters that replace pitchers in the batting order ought to be hit more frequently than pitchers; therefore, we include the season OPS of the batter and a dummy equal to one if a pitcher is hitting. OPS is the sum of the slugging ratio and on-base percentage.[14] A batter with a higher OPS is more likely to generate offense and is therefore more beneficial to hit than a player with a lower OPS. Even when controlling for OPS, pitchers may be less desirable batters to hit relative to teammates, which is why we also include the pitcher dummy. These variables are important in determining whether compositional factors explain the hit batsmen differential.

Though it has not been discussed much in this literature, hit batsmen are largely accidental. Even pitchers who do not mean to brush a player back or pitch on the inside of the plate may accidentally hit batters. The variables in **PQ** control for the pitcher's propensity to hit batters accidentally. Pitchers who are more likely to hit batters mistakenly are also likely to make mistakes in the form of placing hittable balls in the strike zone when compared to other pitchers. These pitchers will generate more offense for opposing teams relative to more skilled pitchers. We use the pitcher's OPS allowed to hitters as a proxy for pitcher skill. Also, pitchers with less control are more likely to issue walks than pitchers with good control. Therefore, we also include the walk rate of the pitcher to control for pitcher quality. Both of these measures ought to be positively related to hit batsmen.

At the heart of this paper is the issue of using the pitcher to retaliate against the other team for hitting batters. This retaliation may be reciprocal, pitcher hitting pitcher, or

[14] Baseball fans may wonder why we chose this measure as opposed to batting average. Though batting average is positively correlated with offensive output, OPS is a better predictor of run production.

pitchers may hit teammates of the pitcher. Trandel (2004) notes that observing retaliation in the data is of fundamental importance to this debate.[15] We have the ability to detect retaliation using the variables in vector **R***.* There are several reasons pitchers retaliate against the other team through plunking, but we will focus on two common reasons that we can capture in the data. If a batter on the current pitcher's team was hit in the previous half-inning, this may provoke a retaliatory plunking of any player on the opposing team. We include a dummy variable equal to one if this occurs. Also, we test specifically for reciprocal retaliation against pitchers. We create a dummy for a pitcher appearing at the plate when a batter was hit in the previous half-inning. If retaliation does occur through pitchers hitting pitchers, this variable should be positive. Also, it is a baseball tradition to hit a batter following a home run, on occasion. We include a dummy equal to one if the previous batter hit a home run.

Finally, we include vector **GS** of many possible game situations affecting the cost-benefit calculation. The score, inning, number of outs, and base runners threatening to score all affect the costs and benefits of hitting a batter at the particular moment. Score differential and whether a team is winning or losing certainly impact the situation. A large absolute score difference means there is a low marginal impact of runs. When the score differential becomes large both teams may more willing to engage in a plunking war than when the game is tight. However, because the costs of retaliation fall as the differential rises, pitchers may be extra careful as the likelihood of retaliation increases. Also, it may be the case that a big loser is more willing to plunk batters. As the score differential grows the losing team may feel the opponent is "rubbing it in," and thus may start plunking the winning team. It is unclear if batters are more likely to be hit

[15] Trandel (2004) states, "While retaliation is widely believed to exist…it is now clear that any attempt to debate the method of retaliation should be postponed. Rather, the attention of interested researchers should be focused on finding a way in which retaliation (to the extent that it exists) can be detected."

early as opposed to later in a game. Plunking early may send a message to players not to "crowd" the inside of the plate throughout the game. However, if retaliation is an *ex post* enforcement mechanism to prevent plunking, later in the game teams have more incentive to renege on tacit agreements not to hit each others' players. The number of outs affects the probability that any base runner will score. A batter put on base with no outs has a better chance of scoring than a batter who reaches first with two outs. Finally, the number and positioning of base runners can also affect the costs and benefits. The specific predictions between differing runner configurations is a bit complicated and largely irrelevant to the discussion here. Generally, the more runners on base and the further the runners are around the bases, the greater the offensive damage a batter can do by putting a ball in play. Therefore, the seven of eight possible runner configurations included (we exclude empty bases for comparison) should be positively related to hit batsmen. We also include year effects to control for unobservable differences in hit batsmen that may change from year to year. Table 1 lists the summary statistics for the variables in both of our datasets.

V. Results and Discussion

Retrosheet provides play-by-play data for both leagues from 1972-1992 and for the 1969 season. From this data we extract two four-season datasets (from the earliest and most recent available years) to estimate the empirical model. For each plate appearance we extract data on the specific game situation and the seasonal performance of the players participating in the event. Initially we focus on the four most recent years of available data, 1989-1992.

Table 2 lists the results of the logit and probit estimations of coefficients and z-statistics. For ease of interpretation we also include the odds ratio (logit) and marginal probability (probit) calculations for each variable. Most of the variables are significant and of the expected sign. Of particular interest, the DH variable is positive and significant when controlling for all other factors, including the compositional batter quality variables. When the DH is in effect pitchers are 15 to 17 percent more likely to hit batters than when the DH is not in effect. Consistent with the composition hypothesis, batter quality is positively related with hit batsmen and pitchers are about 55 percent less likely to be hit than other batters. Also, pitcher quality is negatively related to hit batsmen. Both proxies for lack of pitch control are positively associated with a greater likelihood of hitting a batter; however, pitcher OPS is not statistically significant.

The retaliation variables tell an interesting story. Preceding events likely to provoke retaliation all tend to increase the likelihood that a pitcher hits a batter. A batter who appears at the plate following a home run is 32 percent more likely to be hit than when the preceding batter did not hit a home run. In response to hit batters, pitchers are more likely to hit batters the half-inning after the opposing pitcher hits a current pitcher's teammate. Though the results are not significant for players in general, a pitcher is four times more likely to be hit when an opposing player was hit in the previous half-inning. This variable, which is significant, is very important for the moral hazard hypothesis. We now observe retaliation against pitchers who hit batters. This is a phenomenon previously unidentified in the aggregate data. Though pitchers are hit less frequently than other players on average, pitchers are more likely to be hit after plunking an opposing player. Therefore, pitchers do bear a very real cost to hitting batters, which is necessary to explain a moral hazard response by pitchers.

For the game situation variables, outs and innings seem to be relatively unimportant. Though the absolute score differential is unimportant, the relative score differential from the offense to the defense has a small but significant effect on the incidence of hit batsmen. The base runner composition variables appear to be important and consistent with our predictions.

Over this sample, the AL hit batsmen rate is 26 percent higher than the NL. Controlling for all of the above factors, including batter quality, the DH rule dummy explains approximately 60 percent of this difference. It is now clear that something other than the compositional factors created by the DH is responsible for the increase in hit batsmen differential between leagues. The most plausible explanation for this difference is the moral hazard of pitchers. However, it is possible that some yet unidentified factors unique to the AL may explain the sign and significance of the DH dummy, since the DH occurs only in this league. For example, differences in strike zones, stadium configurations, league traditions, etc. between leagues are competing but less satisfying explanations. To be thorough, we examine a different dataset that includes the 1969 and 1972-1974 seasons.[16] This is the earliest available data in the Retrosheet archives, plus this period includes some observations of the AL without the DH rule. We are slightly less confident in this data than our more recent data for a few reasons. First, the only two seasons available prior to the introduction of the DH are the noncontiguous years 1969 and 1972. Second, the late 1960s and early 1970s were low offense years for baseball. The DH was the last of several rules instituted to "fix" baseball, which was losing fans. The league responded in 1969 by lowering the pitchers mound, shrinking the strike zone to make it more hitter friendly, and adding two new teams to each league. Luckily, the exogenous events other than the DH apply to both leagues; therefore, we feel the data are useful in isolating effect of the DH rule on hit batsmen, given that we acknowledge the deficiencies.

We apply the same model for the previous sample to this data, except the DH variable is now coded zero for some observations in the AL, when the DH was not in effect. Therefore, the DH should now reflect the change in the costs of hitting batters induced by the DH rule itself. The results in Table 3 are quite similar to estimates of the

[16] A few games in the Retrosheet data are missing for the 1972 and 1973 seasons.

more recent data. For this sample, when the DH is in effect pitchers are 11 to 12 percent more likely to hit batters than without the DH rule, *ceteris paribus*. These results differ only slightly in magnitude from our previous estimates. We also observe pitchers as well as all players being hit in retaliation for hit batters in this sample. Most of the other factors are of similar magnitude and size, though some measures are no longer statistically significant. In the years following the DH, AL pitchers hit batters at a rate 14 percent higher than NL pitchers; thus, the DH rule dummy explains approximately 80 percent of the interleague difference in hit batsmen in this sample. Therefore, the results indicate that moral hazard induced by the DH rule, not unobserved league-specific factors, is the likely cause.

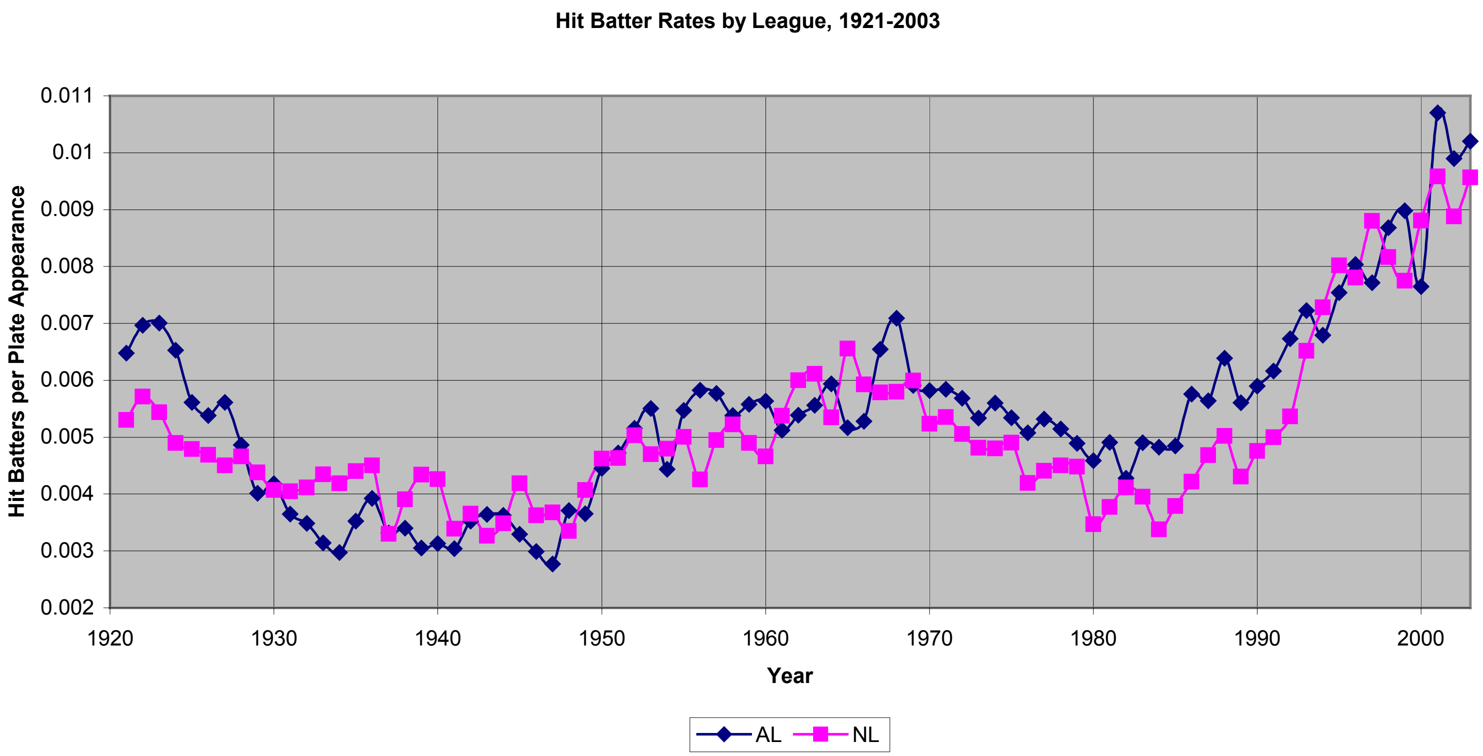
VI. Conclusion

Competing economic theories of hit batsmen predict that DH rule is responsible for raising the rate of hit batsmen in the AL above the NL. As Figure 1 illustrates, for the three decades following the introduction of the DH the AL hit batsmen rate remained higher than the NL rate for all but four years. However, there is less agreement regarding the cause of this difference. While the moral hazard of pitchers in the AL, who do not have to bat and face reciprocal retaliation, is a likely explanation, it is not the only explanation. An alternate hypothesis is that the DH increases the incidence of hit batters by altering the composition of the batting order. Replacing a poor-hitting pitcher with good-hitting DH raises the marginal benefit to hitting an additional batter in the lineup. Additionally, the retaliation threat to pitchers may be too small to be relevant. If retaliation is not likely, there is no cost differential to pitchers between leagues to explain the hit batsmen difference.

Past empirical analyses of the competing hypotheses to explain the hit batsmen differential have relied on aggregate league data. However, the explicit cost-benefit analysis of the pitcher at the time of a pitch is not observable with aggregate data. Using two new unique datasets we have been able to observe the costs and benefits of hitting batters during individual plate appearances. Therefore, we isolate the predicted effects of the competing hypotheses and determine the impact of many relevant factors that influence the pitcher's consumption of risk in hitting a batter every time a hitter steps to the plate. Controlling for variables that proxy batter quality, pitcher quality, retaliation, and game situation we find that the DH rule increases the likelihood that any batter will be hit during a plate appearance between 11 and 17 percent. This explains approximately 60 to 80 percent of the differential in the hit batsmen rate between leagues. Furthermore, we observe that pitchers are more likely to be hit after plunking an opposing player. This indicates that retaliation against pitchers, which is necessary for the moral hazard explanation, does occur. In total, the findings support the original moral hazard explanation of GST even after controlling for batting order composition. Though other factors are important in the economics of plunking, moral hazard plays a nontrivial role in explaining the hit batsmen differential between leagues.

Table 1. Summary Statistics

	1989-1992				1969, 1972-1974			
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Hit By Pitch	0.005528	0.074145	0	1	0.005394	0.073247	0	1
Designated Hitter	0.54037	0.498368	0	1	0.253612	0.435078	0	1
Batter OPS	0.702851	0.126917	0	3	0.687984	0.142806	0	4
Pitcher Hitting	0.029663	0.169656	0	1	0.05493	0.227845	0	1
Pitcher OPS	0.705841	0.089025	0	4	0.690408	0.079808	0	2.5
Pitcher Walk Rate	0.087155	0.027912	0	0.5	0.08878	0.027886	0	1
Batter Hit in Last Half-Inning	0.022312	0.147695	0	1	0.021249	0.144212	0	1
Pitcher Hitting after Batter Hit	0.00056	0.023647	0	1	0.001058	0.032515	0	1
Previous Batter Hit Home Run	0.019982	0.139937	0	1	0.019453	0.138112	0	1
Relative Score (Offense to Defense)	-0.04484	2.959261	-20	20	-0.04525	2.924196	-19	19
Absolute Relative Score	2.017926	2.164995	0	20	1.96396	2.166984	0	19
Inning	5.035247	2.6935	1	22	5.067495	2.735999	1	25
Outs	0.977866	0.81623	0	2	0.977184	0.816156	0	2
Man on First	0.179398	0.383686	0	1	0.188647	0.391228	0	1
Man on Second	0.08862	0.284194	0	1	0.083079	0.276003	0	1
Man on Third	0.031418	0.174445	0	1	0.027007	0.162104	0	1
Men on First and Second	0.068588	0.252753	0	1	0.071399	0.25749	0	1
Men on First and Third	0.032199	0.176528	0	1	0.031505	0.174679	0	1
Men on Second and Third	0.022048	0.14684	0	1	0.020134	0.140458	0	1
Bases Loaded	0.022591	0.148594	0	1	0.02336	0.151045	0	1



Source: *The Lahman Baseball Database*, Version 5.1 (2004)

Table 2. Estimates of Hit Batsmen, 1989-1992

Category	Variable	Logit		Probit	
Moral Hazard	Designated Hitter Dummy	*Coefficient*	0.162214	*Coefficient*	0.056729
		Odds Ratio	1.176112	*Marg. Prob*	0.000864
		z-statistic	4.63**	*z-statistic*	4.65**
Batter Quality	Batter OPS	*Coefficient*	0.614065	*Coefficient*	0.215086
		Odds Ratio	1.847929	*Marg. Prob*	0.003293
		z-statistic	3.93**	*z-statistic*	3.95**
	Pitcher Hitting Dummy	*Coefficient*	-0.81375	*Coefficient*	-0.26391
		Odds Ratio	0.443191	*Marg. Prob*	-0.00297
		z-statistic	-4.52**	*z-statistic*	-4.6**
Pitcher Quality	Pitcher OPS	*Coefficient*	0.235353	*Coefficient*	0.086929
		Odds Ratio	1.265355	*Marg. Prob*	0.001331
		z-statistic	1.24	*z-statistic*	1.3
	Pitcher Walk Rate	*Coefficient*	2.893827	*Coefficient*	1.033456
		Odds Ratio	18.06231	*Marg. Prob*	0.015822
		z-statistic	4.89**	*z-statistic*	4.92**
Retaliation	Hit Batter in Previous Half-Inning	*Coefficient*	0.073132	*Coefficient*	0.023006
		Odds Ratio	1.075873	*Marg. Prob*	0.000362
		z-statistic	0.65	*z-statistic*	0.58
	Pitcher Hitting after Batter Hit in Previous Half-Inning	*Coefficient*	1.447219	*Coefficient*	0.485283
		Odds Ratio	4.251276	*Marg. Prob*	0.013967
		z-statistic	2.36**	*z-statistic*	2.18*
	Previous Batter Hit Home Run	*Coefficient*	0.283919	*Coefficient*	0.099798
		Odds Ratio	1.328325	*Marg. Prob*	0.001729
		z-statistic	2.68**	*z-statistic*	2.62**
Game Situation	Relative Score Differential of Offense to Defense	*Coefficient*	0.042483	*Coefficient*	0.01472
		Odds Ratio	1.043399	*Marg. Prob*	0.000225
		z-statistic	7.02**	*z-statistic*	7.05**
	Absolute Relative Score	*Coefficient*	-0.0117	*Coefficient*	-0.00368
		Odds Ratio	0.988371	*Marg. Prob*	-5.6E-05
		z-statistic	-1.38	*z-statistic*	-1.26
	Inning	*Coefficient*	-0.01545	*Coefficient*	-0.0055
		Odds Ratio	0.984668	*Marg. Prob*	-8.4E-05
		z-statistic	-2.31*	*z-statistic*	-2.36**

	Outs	*Coefficient*	-0.01707	*Coefficient*	-0.00581
		Odds Ratio	0.983078	*Marg. Prob*	-8.9E-05
		z-statistic	-0.8	*z-statistic*	-0.78
	Man on First	*Coefficient*	0.048279	*Coefficient*	0.016536
		Odds Ratio	1.049463	*Marg. Prob*	0.000257
		z-statistic	1.03	*z-statistic*	1.01
	Man on Second	*Coefficient*	0.075331	*Coefficient*	0.026281
		Odds Ratio	1.078241	*Marg. Prob*	0.000414
		z-statistic	1.23	*z-statistic*	1.23
Game Situation (continued)	Man on Third	*Coefficient*	0.278385	*Coefficient*	0.098077
		Odds Ratio	1.320995	*Marg. Prob*	0.00169
		z-statistic	3.14**	*z-statistic*	3.1**
	Men on First and Second	*Coefficient*	0.135049	*Coefficient*	0.047716
		Odds Ratio	1.144593	*Marg. Prob*	0.00077
		z-statistic	2.02*	*z-statistic*	2.03*
	Men on First and Third	*Coefficient*	0.197402	*Coefficient*	0.068936
		Odds Ratio	1.218234	*Marg. Prob*	0.001147
		z-statistic	2.19*	*z-statistic*	2.15*
	Men on Second and Third	*Coefficient*	0.251846	*Coefficient*	0.08975
		Odds Ratio	1.286398	*Marg. Prob*	0.001534
		z-statistic	2.4**	*z-statistic*	2.4**
	Bases Loaded	*Coefficient*	0.132466	*Coefficient*	0.045277
		Odds Ratio	1.14164	*Marg. Prob*	0.000733
		z-statistic	1.21	*z-statistic*	1.17
Year Dummies	Year 1990	*Coefficient*	0.050283	*Coefficient*	0.017926
		Odds Ratio	1.051568	*Marg. Prob*	0.000278
		z-statistic	1.02	*z-statistic*	1.05
	Year 1991	*Coefficient*	0.096062	*Coefficient*	0.033389
		Odds Ratio	1.100827	*Marg. Prob*	0.000522
		z-statistic	1.97*	*z-statistic*	1.97*
	Year 1992	*Coefficient*	0.189908	*Coefficient*	0.066534
		Odds Ratio	1.209139	*Marg. Prob*	0.001064
		z-statistic	3.97**	*z-statistic*	3.99**
Goodness-Of-Fit	Observations		641640		641640
	Log-likelihood		-21841.72		-21841.24
	LR- χ^2		264.99		265.96
	Prob. > (Pearson χ^2)		0.54		0.50

Table 3. Estimates of Hit Batsmen, 1969 and 1972-1974

Category	Variable	Logit		Probit	
Moral Hazard	Designated Hitter Dummy	*Coefficient*	0.116442	*Coefficient*	0.039419
		Odds Ratio	1.123492	*Marg. Prob*	0.000596
		z-statistic	2.25*	*z-statistic*	2.19*
Batter Quality	Batter OPS	*Coefficient*	1.415542	*Coefficient*	0.509973
		Odds Ratio	4.118717	*Marg. Prob*	0.007512
		z-statistic	10.07**	*z-statistic*	9.87**
	Pitcher Hitting Dummy	*Coefficient*	-0.45747	*Coefficient*	-0.14043
		Odds Ratio	0.632884	*Marg. Prob*	-0.00177
		z-statistic	-3.68**	*z-statistic*	-3.41**
Pitcher Quality	Pitcher OPS	*Coefficient*	0.708161	*Coefficient*	0.266846
		Odds Ratio	2.030253	*Marg. Prob*	0.003931
		z-statistic	3.2*	*z-statistic*	3.33*
	Pitcher Walk Rate	*Coefficient*	3.173758	*Coefficient*	1.150442
		Odds Ratio	23.89712	*Marg. Prob*	0.016947
		z-statistic	5.16**	*z-statistic*	5.16**
Retaliation	Hit Batter in Previous Half-Inning	*Coefficient*	0.334874	*Coefficient*	0.120795
		Odds Ratio	1.397764	*Marg. Prob*	0.002068
		z-statistic	3.04**	*z-statistic*	3.06**
	Pitcher Hitting after Batter Hit in Previous Half-Inning	*Coefficient*	1.630977	*Coefficient*	0.581011
		Odds Ratio	5.108862	*Marg. Prob*	0.018297
		z-statistic	4.57**	*z-statistic*	4.24**
	Previous Batter Hit Home Run	*Coefficient*	0.172655	*Coefficient*	0.06083
		Odds Ratio	1.188456	*Marg. Prob*	0.000966
		z-statistic	1.51	*z-statistic*	1.48
Game Situation	Relative Score Differential of Offense to Defense	*Coefficient*	0.046296	*Coefficient*	0.016223
		Odds Ratio	1.047385	*Marg. Prob*	0.000239
		z-statistic	7.22**	*z-statistic*	7.31**
	Absolute Relative Score	*Coefficient*	0.006134	*Coefficient*	0.0027
		Odds Ratio	1.006153	*Marg. Prob*	3.98E-05
		z-statistic	0.7	*z-statistic*	0.89
	Inning	*Coefficient*	-0.01564	*Coefficient*	-0.00546
		Odds Ratio	0.984482	*Marg. Prob*	-8.1E-05
		z-statistic	-2.26*	*z-statistic*	-2.27*

Game Situation (continued)	Outs	*Coefficient*	0.020185	*Coefficient*	0.006556
		Odds Ratio	1.02039	*Marg. Prob*	9.66E-05
		z-statistic	0.89	*z-statistic*	0.83
	Man on First	*Coefficient*	-0.06239	*Coefficient*	-0.02133
		Odds Ratio	0.939517	*Marg. Prob*	-0.00031
		z-statistic	-1.25	*z-statistic*	-1.24
	Man on Second	*Coefficient*	0.144596	*Coefficient*	0.050283
		Odds Ratio	1.155572	*Marg. Prob*	0.000782
		z-statistic	2.29*	*z-statistic*	2.25*
	Man on Third	*Coefficient*	0.09539	*Coefficient*	0.034457
		Odds Ratio	1.100088	*Marg. Prob*	0.000529
		z-statistic	0.91	*z-statistic*	0.93
	Men on First and Second	*Coefficient*	0.072292	*Coefficient*	0.024472
		Odds Ratio	1.074969	*Marg. Prob*	0.00037
		z-statistic	1.04	*z-statistic*	1
	Men on First and Third	*Coefficient*	-0.09798	*Coefficient*	-0.03147
		Odds Ratio	0.906663	*Marg. Prob*	-0.00045
		z-statistic	-0.92	*z-statistic*	-0.85
	Men on Second and Third	*Coefficient*	0.05883	*Coefficient*	0.021674
		Odds Ratio	1.060595	*Marg. Prob*	0.000328
		z-statistic	0.48	*z-statistic*	0.5
	Bases Loaded	*Coefficient*	-0.05206	*Coefficient*	-0.0202
		Odds Ratio	0.949273	*Marg. Prob*	-0.00029
		z-statistic	-0.43	*z-statistic*	-0.48
Year Dummies	Year 1969	*Coefficient*	0.191714	*Coefficient*	0.066753
		Odds Ratio	1.211324	*Marg. Prob*	0.001027
		z-statistic	3.41**	*z-statistic*	3.4**
	Year 1972	*Coefficient*	0.168274	*Coefficient*	0.05901
		Odds Ratio	1.18326	*Marg. Prob*	0.000905
		z-statistic	2.89**	*z-statistic*	2.91**
	Year 1973	*Coefficient*	-0.05735	*Coefficient*	-0.02012
		Odds Ratio	0.944265	*Marg. Prob*	-0.00029
		z-statistic	-1.11	*z-statistic*	-1.13
Goodness-Of-Fit	Observations		584886		584886
	Log-likelihood		-19436.32		-19434.23
	LR- χ^2		373.84		378.01
	Prob. > (Pearson χ^2)		0.98		0.89

Notes for Tables

Constants not reported.

* = statistical significance at the 5 percent level.

** = statistical significance at the 1 percent level.

References

Bradbury, J.C and Drinen, D.J. (2004). “The Designated Hitter, Moral Hazard, and Hit Batters: New Evidence from Game-Level Data.” Manuscript. The University of the South.

Duggan, M. and Levitt, S.D. (2002). “Winning Isn’t Everything: Corruption in Sumo Wrestling.” *American Economic Review*, 92: 1594-1605.

Goff, B.L., McCormick, R.E., and Tollison, R.D. (2002). “Racial Integration as an Innovation: Empirical Evidence from Sports Leagues.” *American Economic Review*, 92: 16-26.

Goff, B.L.; Shughart, W.F.; and Tollison, R.D. (1997). “Batter Up! Moral Hazard and the Effects of the Designated Hitter Rule on Hit Batsmen.” *Economic Inquiry*, 35: 555-561.

Goff, B.L.; Shughart, W.F.; and Tollison, R.D. (1998). “Moral Hazard and the Effects of the Designated Hitter Rule Revisited.” *Economic Inquiry*, 36: 688-692.

Gould, S.J. (1996). *Full House: The Spread of Excellence from Plato to Darwin*. New York: Harmony Books.

Heckelman, J.C. and Yates, A.J. (2003). “And a Hockey Game Broke Out: Crime and Punishment in the NHL.” *Economic Inquiry*, 41: 705-712.

James, B. (1985). “A History of the Beanball.” *The Bill James Baseball Abstract*, New York: Ballantine Books, 131-140.

Kawaura, A. and La Croix, S.J. (2002). “Baseball in Japan and North America: Same Game, Same Rules, Same Results?” Manuscript. University of Hawaii.

The Lahman Baseball Database. (2004), Version 5.1. (www.baseball1.com).

Levitt, S.D. (1998). “The Hazards of Moral Hazard: Comment on Goff, Shughart, and Tollison.” *Economic Inquiry*, 36: 685-687.

McCormick, R.E., and Tollison, R.D. (1984). “Crime on the Court.” *Journal of Political Economy*, 92: 223-235.

Morgan, J. (2002). “Attention Estes: Throw at Clemens.” *ESPN.com*, June 14, 2002.

N.A.P.B.L.Umpire Manual (1996). Chicago: Triumph Books.

Pappas, Doug (1997). ”Terms of Major League Expansion.” *Outside the Lines*, 3(Fall).

Peltzman, S. (1975). “The Effects of Automobile Safety Regulation.” *Journal of Political Economy*, 83: 677-725.

Retrosheet. (www.retrosheet.org/game.htm).

Rosenthal, K. (2002). “Mets Get Shot with Mighty Clemens at the Bat.” *The Sporting News*, June 13, 2002.

Rupert, M.E. and Stephenson, E.F. (2002). “On Moral Hazard and Hit Batsmen: A Comment on Goff *et. al*. and Trandel *et. al*.” Manuscript. Berry College.

Schmidt, M.B. and Berri, D.J. (2003). “On the Evolution of Competitive Balance: The Impact of Increasing Global Search.” *Economic Inquiry*, 41: 692-704.

Sobel, R.S.and Nesbitt T.M. (2004). “Automobile Safety Regulation and the Incentive to Drive Recklessly: Evidence from NASCAR.” Manuscript. West Virginia University.

Trandel, G.A. (2004). “Hit By Pitches: Moral Hazard, Cost-Benefit, Retaliation, or Lack of Evidence?” *Journal of Sports Economics*, 5: 87-92

Trandel, G.A.; White, L.H.; and Klein, P.G. (1998). “The Effect of the Designated Hitter Rule on Hit Batsmen: Pitcher’s Moral Hazard or the Team’s Cost-Benefit Calculation.” *Economic Inquiry*, 36: 679-684.

Turocy, T.L. (2003). “Offensive Performance, Omitted Variables, and the Value of Speed in Baseball.” Manuscript. Texas A&M University.