Research Notes: The Designated Hitter, Moral Hazard, and Hit Batters: New Evidence From Game-Level Data
John Charles Bradbury and Douglas Drinen
Journal of Sports Economics 2006; 7; 319
DOI: 10.1177/1527002504273391

The online version of this article can be found at: http://jse.sagepub.com/cgi/content/abstract/7/3/319
The Designated Hitter, Moral Hazard, and Hit Batters
New Evidence From Game-Level Data

JOHN CHARLES BRADBURY
DOUGLAS DRINEN
Sewanee: The University of the South

The rate of hit batters in the American League (AL) has exceeded the National League (NL) by an average rate of 15% per season since the introduction of the designated hitter (DH) to the AL. This difference may be the result of moral hazard by AL pitchers; however, past studies have been unable to differentiate between competing explanations. Using game-level statistics from the entire 31-year history of the DH, we find the DH to be positively correlated with hit batsmen while controlling for many factors. Results indicate that the difference in hit batters between leagues is real and that moral hazard explains about one half of the difference. Observations from interleague play further support the notion that the results are not an artifact of unidentified league-specific factors.

Keywords: moral hazard; baseball; hit batters

In 1973, the American League (AL) of Major League Baseball (MLB) introduced the designated hitter rule, which allows a designated hitter (DH) to bat in place of the pitcher. Although the intention of the rule change was to increase offense in the game, it also had the unintended consequence of increasing the frequency of hit batters in the AL. The National League (NL), which continues to disallow the DH, has averaged a hit batter rate 15% lower than the AL since the introduction of the DH. As Figure 1 illustrates, the AL hit batter rate has exceeded the NL rate for all but four seasons of the 31-season history of the DH. This rule change has been of particular interest to economists because it is consistent with rational response predicted by the law of demand. Relative to their NL counterparts, AL pitchers face a

AUTHORS’ NOTE: Thanks to David Smith and the volunteers of the Retrosheet organization for compiling and sharing the data.
lower price for hitting batters—because AL pitchers do not have to bat where they might face retaliation—therefore, they ought to hit batters with more frequency. This does not mean that pitchers necessarily hit batters on purpose. AL pitchers may engage in activities, such as pitching high and inside, that are more likely to result in hitting the batter. Unlike NL pitchers, AL pitchers are not monitored by the opposing team’s pitcher. The teammates and manager of the pitcher, in addition to the opposing batters, may suffer as a consequence, which makes this a problem of moral hazard.1

However, the persistent hit-batsmen differential between leagues is not sufficient to prove the DH creates a moral hazard response from pitchers. Although removing the pitcher’s obligation to bat certainly prevents retaliation by the opposition, it also removes a very weak hitter from the lineup. Almost always, pitchers are poor hitters who represent easy outs. The punishment for hitting a batter is to advance the hit batter to first base, making the opposing team extra careful not to hit pitchers. Hence, the hit batsmen differential between the leagues could also be the consequence of changing the composition of the batting orders in the AL by adding a batter whom pitchers are more willing to risk hitting. In addition, the hit batter rate difference between leagues began to narrow in the mid-1990s, calling into question whether the sustained difference of the preceding 20 years was real or simply a statistical run regressing to the mean.

Several studies, which we discuss in the following section, debate validity of the competing hypotheses. Although informative, most of these studies use yearly aggregate statistics that are not ideal for distinguishing between the competing

![Figure 1: Hit Batter Rates by League, 1921-2003](http://jse.sagepub.com)

*Figure 1:* Hit Batter Rates by League, 1921-2003

*Source: The Lahman Baseball Database (Version 5.1) (2004).*
hypotheses. Hit batters are rare events in baseball. A retaliatory flurry of beanball pitches in a single game may be lost in a 162-game season; however, it may have a real impact that persists into the future by establishing a credible threat of retaliation that opposing teams heed. In addition, the costs and benefits of hitting a batter in a particular game situation may deter retaliation in one instance but not in another. More specific game information is necessary to fully identify the causes of hit batters. In this article, we use game-level data to examine the impact of the DH rule on specific games for the entire 31-year history of the DH. In addition, this data provides the opportunity to examine interleague games, where NL and AL teams play each other with and without the DH rule in effect. Using a sample of more than 130,000 team games over the history of the DH rule, we find that in DH games batters were about 8% more likely to be hit than in non-DH games. This explains approximately one half of the difference between the leagues during the DH era. In addition, evidence from more than 3,000 interleague games indicates an even stronger moral hazard response to the DH, which is consistent with league-specific factors other that the DH narrowing the difference in hit batsmen rates between leagues. Furthermore, we find evidence of teams hitting batters as a weapon of retaliation.

STATE OF THE DEBATE

In the initial study of the effect of the DH on hit batters, Goff, Shughart, and Tollison (1997) find the hit batsmen differential between leagues to be correlated with the introduction of the DH, which is consistent with the moral hazard hypothesis. However, several follow-up studies view the moral hazard hypothesis with suspicion. Trandel, White, and Klein (1998) and Levitt (1998) point out that the post-DH rise of hit batmen in the AL does not necessitate moral hazard by pitchers. Because pitchers are not good hitters, opposing teams are going to be especially careful not to hit pitchers and put a high-percentage out on base. Therefore, the difference between leagues is consistent with the alternate batting composition hypothesis. In addition, Levitt (1998) finds that retaliation against pitchers who hit batters is much too rare to affect a pitcher’s decision calculus. Similarly, Trandel (2004) finds no evidence, during the course of a season, that opponents hit batters on teams whose pitchers hit batters more frequently. To complicate matters, about the same time these articles were written the difference in hit batsmen rates between the leagues began to fall. From 1994-2000, the NL rate exceeded the AL rate 4 times. Thus, it appeared the moral hazard hypothesis might not be correct.

However, one problem from which these studies suffer is their reliance on aggregate yearly statistics. Goff et al. (1997) and Trandel et al. (1998) use seasonal league differences in hit batsmen rates between leagues. Levitt (1998) examines seasonal correlations between pitchers who hit batters and how often pitchers are hit themselves. Trandel (2004) uses data aggregated by season and team. These aggregations are problematic for the study of hit batters in baseball. Although these
studies are informative, disentangling the competing hypotheses requires micro-
level data. Hit batters in baseball are quite rare events—in 2003, about 1% of all
plate appearances resulted in a hit batter—and the yearly statistics can mask the
costs and benefits involved in hitting batters at any point in time. In addition, hit bat-
ters are largely accidental, which adds noise to aggregate statistics. In an effort to
solve this problem, Bradbury and Drinen (2004) utilize previously unavailable
play-by-play data to isolate the many factors that may cause a pitcher to hit a batter.
Because this data in Bradbury and Drinen (2004) is play specific, the study controls
for the costs and benefits of hitting batters during individual plate appearances.
While controlling for the quality of the batter, the quality of the pitcher, and the
game situation, the authors find the DH to be associated with more hit batters. In
addition, the study finds strong evidence of retaliation against pitchers who hit bat-
ters, which must exist for the moral hazard theory to be true. For the two time peri-
ods studied, the findings indicate that moral hazard explains between 60% and 80%
of the difference in hit batsmen between leagues.

One problem that the previous literature has not been able to solve is the issue of
what happened in the 1990s. None of these studies determine if the shrinking dif-
ference in hit batsmen between leagues is evidence against the moral hazard or bat-
ter composition theories. Although such a long-run difference in hit batter rates
seems highly unlikely by chance, it is possible. There are two exogenous events
that may explain the narrowing of the hit batsmen differential consistent with the
moral hazard and batter composition hypotheses even though the rules governing
the DH remained in place. First, the league expansion of the NL in 1993 employed
expansion rules that affected the leagues asymmetrically, causing the NL to take on
more of the burden of new fringe players who are more apt to hit batters and be hit
by pitches. The expansion explanation is consistent with the overall rise in hit bat-
ters in both leagues, and the greater rise in the NL compared to the AL. Second, in
an effort to limit retaliatory hit batters, MLB instituted the so-called double-warning
rule in 1994, which authorizes umpires to warn both teams if the umpire deems a
plunking to be intentional. The new penalty for retaliation for a previous plunking
is the immediate ejection of the offending pitcher and manager, plus monetary fines
with the possibility of suspension. This rule significantly raised the cost of oppos-
ing teams retaliating for hitting batters, thereby lessening the retaliatory fear among
NL pitchers. Consistent with the moral hazard hypothesis, NL pitchers now bear a
lower cost for hitting batters and therefore ought to be expected to behave more like
their counterparts in the AL. When these two factors are included in the analysis,
the narrowing of hit batsmen rates is understandable; however, further testing is
needed.

In this article, we contribute to this literature by using game-level data to analyze
the hit batters during the entire DH era. Although the game-level data is not as spe-
cific as Bradbury and Drinen (2004), it is the most disaggregated data that extends
to the present and is available to the public. Another advantage of the recent
game-level data is that it includes observations from interleague play, which began
in 1997. Prior to this time, AL and NL teams only played teams within their own leagues during the regular season. After 1997, teams played a portion of their schedule versus teams in the other league. Because the leagues differ in regard to the DH rule, the league of the home team determines whether the DH is allowed. For games in AL parks, the DH is permitted; however, in NL parks it is forbidden. This change provides a unique opportunity to view NL and AL teams playing the game with and without the DH.

EMPIRICAL ANALYSIS

A data-gathering project known as Retrosheet has compiled game-specific data for each season for almost all games from the early 1970s until the present, which roughly coincides with the history of the DH in MLB. Using Retrosheet “game logs,” we generate a data set of team games from 1973-2003.7 At this level we control for game-specific situations that likely influence the amount of hit batters in the game. Using the empirical model of Bradbury and Drinen (2004) as a guide, we estimate the Equation 1 using a Poisson regression.

\[
HBP_g = \alpha + \beta_1DH + \beta_2BQ + \gamma PQ + \phi R + \psi GS + \eta YR + \varepsilon_g
\]  

\(HBP\) is the count of batters hit in each team game. Subscript \(g\) represents a team game, such that each of the teams playing in a game represent separate observations. The \(\alpha\) is a constant, and the \(\varepsilon\) is the random error term.

\(DH\) is a dummy variable equal to one when the game employs the DH rule and zero otherwise. This differs from past studies because the DH is in effect for some observations of NL teams and not in effect for some AL teams during interleague games. If the DH causes moral hazard, then, after controlling for the factors that follow, this variable ought to be positively correlated with hit batters in games.

To control for the influence of offensive prowess on hit batters, Bradbury and Drinen (2004) use a vector of batter quality (BQ) variables that includes controls for factors associated with individual batters being hit as a result of hitting ability. The batter composition hypothesis predicts that better batters are more beneficial to hit than worse batters. Because we use game data, not individual players, we only need the seasonal runs scored per game for the team to measure team-hitting ability. We expect this variable to be positive as predicted by the batter composition hypothesis.

A point that we have not stressed in this article is that pitchers often hit batters accidentally for reasons other that moral hazard and batter quality. We include two variables in vector PQ to control for pitcher quality in preventing pitchers from hitting batters accidentally. The average runs allowed per game for the season of analysis by the pitching team proxies the quality of pitching. Pitchers who hit batters frequently but unintentionally are also more likely to commit errors by pitching more hittable balls to the opposition, which should result in more offense for the
opposing teams. In addition, some pitchers may have less ability to keep the ball in the strike zone but still be able to get batters out. These pitchers may not allow many runs to score, while still hitting batters more frequently than the average pitcher. To capture the effect of a pitching team having such so-called wild pitchers we use the seasonal average walks per game. Both of these variables ought to be positively related to hit batsmen.

Retaliation is of central importance to the moral hazard hypothesis. While we cannot identify retaliation against pitchers with game-level data, we can identify retaliation in general as a motive for hitting batters. Vector R includes two variables that might provoke pitchers to hit batters as retaliation. The most obvious motive for retaliation is hitting the other team’s batters; therefore, we include the number of batters hit by the team’s pitchers in that game. In addition, hitting home runs can incite a pitcher to hit a batter. Whether it is because of players “showing up” the pitcher or just plain spite, it is a cultural phenomenon in baseball to plunk batters following a home run, on occasion. Bradbury and Drinen (2004) find support for both types of retaliation.

The variables in vector GS include two game-specific variables that may influence hit batters: the relative score difference and the absolute difference in score. Relative score is the score differential of the team of analysis minus the score of the opposing team. The score differential increases as the winning team scores more runs than the losing team. This variable captures the “sore loser” impact on hit batsmen. A team that is winning by many runs may incite the losing team to hit batters if the loser feels the winner is “rubbing it in.” The absolute score differential is the absolute value of the relative score differential, which captures the declining value of runs in the game regardless of the win-loss outcome. As the score difference grows, the cost of putting another runner on base by hitting a batter falls. Pitchers for the winner and the loser may become less careful as the consequence of the marginal hit batter becoming less important to the outcome of the game. However, as the win-loss consequences diminish, the likelihood of retaliation for hitting batters increases. Hence, pitchers may be even more careful than when the game is close. Finally, we include a vector of year dummies (YR) to control for season-specific factors influencing the rate of hit batters. Table 1 lists summary statistics for all of the variables.

First, we examine all the games of the DH era together. Table 2 lists the Poisson regression estimates for the entire sample. The deviance statistic indicates that the data are not overdispersed, therefore the Poisson regression is appropriate. We report the estimates as incident rate ratios for ease of interpretation, and we use several different model specifications to ensure robustness of the estimates. The results support the moral hazard and the batter composition hypotheses. A game with the DH is associated with between 5% and 12% more hit batters than a non-DH game, depending on which factors we control for in the regression. Consistent with the batter composition hypothesis, an additional run scored per game increases the likelihood of a hit batter by between 2% and 9%. In the full specification (Model 1),
where we control for all relevant factors, the DH increases the likelihood of a hit batter by nearly 8%. During the entire history of the DH, the AL rate has exceeded the NL rate by 15%; therefore, roughly one half of this difference is explained by

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M  SD  Min  Max</td>
<td>M  SD  Min  Max</td>
</tr>
<tr>
<td>Designated hitter</td>
<td>0.5117 0.4999 0 1</td>
<td>0.4710 0.4992 0 1</td>
</tr>
<tr>
<td>Runs scored per game</td>
<td>4.4844 0.5545 3.1 6.23</td>
<td>4.8443 0.5253 3.54 6.23</td>
</tr>
<tr>
<td>Runs allowed per game</td>
<td>4.4845 0.5699 3.01 6.81</td>
<td>4.8443 0.5588 3.43 6.35</td>
</tr>
<tr>
<td>Walks allowed per game</td>
<td>3.3347 0.4116 2.3 4.84</td>
<td>3.4499 0.4105 2.3 4.55</td>
</tr>
<tr>
<td>Batters hit</td>
<td>0.2410 0.5081 0 5</td>
<td>0.3466 0.6065 0 5</td>
</tr>
<tr>
<td>Home runs</td>
<td>0.8858 1.0039 0 10</td>
<td>1.0882 1.1041 0 9</td>
</tr>
<tr>
<td>Runs ahead and/or behind</td>
<td>0 4.2129 –22 22</td>
<td>0 4.4747 –22 22</td>
</tr>
<tr>
<td>Absolute score</td>
<td>3.3381 2.5702 0 22</td>
<td>3.5494 2.7248 0 22</td>
</tr>
</tbody>
</table>

NOTE: Min = minimum; Max = maximum.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Designated hitter</td>
<td>1.077 (5.75)***</td>
<td>1.049 (3.70)***</td>
<td>1.05 (3.77)***</td>
<td>1.094 (7.31)***</td>
<td>1.122 (9.98)***</td>
</tr>
<tr>
<td>Runs scored per game</td>
<td>1.017 (1.28)</td>
<td>1.059 (4.24)***</td>
<td>1.072 (5.17)***</td>
<td>1.085 (6.08)***</td>
<td></td>
</tr>
<tr>
<td>Runs allowed per game</td>
<td>1.077 (5.06)***</td>
<td>1.121 (7.81)***</td>
<td>1.132 (8.50)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walks allowed per game</td>
<td>1.115 (6.21)***</td>
<td>1.126 (6.72)***</td>
<td>1.122 (6.52)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batters hit</td>
<td>1.148 (13.29)***</td>
<td>1.098 (8.95)***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home runs</td>
<td>0.973 (4.59)***</td>
<td>1.05 (8.90)***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Runs ahead and/or behind</td>
<td>1.054 (35.45)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absolute score</td>
<td>1.01 (4.45)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>131342</td>
<td>131342</td>
<td>131342</td>
<td>131342</td>
<td>131342</td>
</tr>
<tr>
<td>Log likelihood ratio $\chi^2$</td>
<td>–77519.52</td>
<td>–78261.27</td>
<td>–78342.17</td>
<td>–78466.38</td>
<td>–78485.95</td>
</tr>
<tr>
<td>Likelihood ratio $\chi^2$</td>
<td>4859.19***</td>
<td>3375.67***</td>
<td>3213.87***</td>
<td>2965.47***</td>
<td>2926.32***</td>
</tr>
<tr>
<td>Wald $\chi^2$</td>
<td>5134.67***</td>
<td>3399.18***</td>
<td>3189.02***</td>
<td>2909.13***</td>
<td>2854.48***</td>
</tr>
</tbody>
</table>

NOTE: Incidence rate ratios with robust $z$ statistics in parentheses. Constants and year effects not reported.

$p < .10$. **$p < .05$. ***$p < .001$. 

where we control for all relevant factors, the DH increases the likelihood of a hit batter by nearly 8%. During the entire history of the DH, the AL rate has exceeded the NL rate by 15%; therefore, roughly one half of this difference is explained by
Almost all of the variables are statistically significant and have the predicted sign across the different specifications. The main exceptions are home runs and runs scored per game in the full model. The former result is not surprising because retaliation for home runs occurs only occasionally. The lack of significance for runs scored in the full model is likely because of multicollinearity with the game situation variables.

The hit-batter retaliation variable is particularly interesting because some past research in this area has been unable to identify any retaliation for hitting batters. Each batter that a team hits in a game increases the team’s hit batters by 10% to 15%. Although we cannot identify pitchers who hit batters being hit, as Bradbury and Drinen (2004) do with play-by-play data, it now seems likely that teams do retaliate for hitting batters. Thus, pitchers may realize that shirking on the mound can result in retaliation against themselves.

It is possible that the incidence rate of hit batters in the sample may be picking up unidentified league-specific differences that influence hit batsmen. Although we include some observations of games where both leagues operate under the rules of their sister league, we want to be certain that any league effects are not masked by the volume of DH–AL and non-DH–NL games in the sample. Therefore, we estimate the full model focusing only on the interleague era, where both leagues play each other and operate under different DH rules. Table 3 lists the Poisson regression estimates (again listed as incidence rate ratios) for all games during that era of interleague play (1997-2003), noninterleague games, and interleague games. The results confirm that the DH reflects something other than league-specific factors, and they also support the contention that some institutional changes in the NL may be hiding the true moral hazard effects of the DH on hit batters. For the entire sample, the DH is associated with just more than 4% more hit batters a game than without the DH, controlling for relevant factors. During this time period, the AL hit batter rate exceeded the NL rate by an average of slightly less than 4%. The estimate indicates that the DH is responsible for all of, if not more than, the actual difference in hit batters between leagues during the interleague era. Thus, other factors that asymmetrically affect the NL are the likely cause of the narrowing of hit batsmen between leagues. When we break the sample into interleague and noninterleague games the results are even more informative. In noninterleague games, having a DH is associated with just less than 4% more hit batters than non-DH interleague games; however, the results are not statistically significant. In contrast, in interleague games the DH is associated with an 11% increase in the incidence of hit batsmen. These results are statistically significant at about the 6% level. This is important because the league-specific factors associated with the DH rule are no longer correlated with the DH rule for these games. When NL and AL teams have a DH, both are more apt to hit batters at a higher rate than without the DH.
DISCUSSION AND CONCLUSION

The debate about the cause of the difference in hit batter rates between the AL and NL has proven fruitful. Early attempts to test economic hypotheses using this natural experiment yielded several competing hypotheses that required further investigation. Using game-level data, we have been able to control for game-specific factors that allow us to differentiate between alternate hypotheses. The results indicate that the moral hazard hypothesis proposed by Goff et al. (1997) is largely correct. As an alternate explanation for the difference, subsequent articles suggested that the difference in hit batsmen could be the result of the offensive potential of AL lineups without any role for moral hazard. This batter composition hypothesis is correct—better batters are positively associated with hit batters—however, it is not sufficient to explain the entire impact of the DH on hit batters. After accounting for offensive power and several other potential factors, having the DH is associated with a greater incidence of hit batters at a rate that explains about one half of the difference between the two leagues. In addition, we find further evidence that teams hit batters as a form of retaliation. Coupled with the findings of Bradbury and Drinen (2004), which find that pitchers are more likely to be hit the inning after hitting a batter on the opposing team, the retaliation required by the moral hazard hypothesis does exist. In addition, the near convergence of hit batsmen rates in the mid-1990s does not appear to be evidence against moral hazard. The exogenous shocks of expansion and the double-warning rule seem to be more plausible explanations for the narrowing. In summary, we believe the findings support the existence of moral hazard among pitchers as a result of the DH.

TABLE 3: Determinants of Hit Batters in the Interleague Era, 1997-2003

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Intraleague</th>
<th>Interleague</th>
</tr>
</thead>
<tbody>
<tr>
<td>Designated hitter</td>
<td>1.044 (2.08)**</td>
<td>1.037 (1.58)</td>
<td>1.113 (1.83)*</td>
</tr>
<tr>
<td>Runs scored</td>
<td>0.989 (0.57)</td>
<td>0.992 (0.37)</td>
<td>0.986 (0.24)</td>
</tr>
<tr>
<td>Runs allowed</td>
<td>1.037 (1.69)*</td>
<td>1.038 (1.62)</td>
<td>1.051 (0.81)</td>
</tr>
<tr>
<td>Walks allowed</td>
<td>1.154 (4.80)***</td>
<td>1.165 (4.79)***</td>
<td>1.033 (0.37)</td>
</tr>
<tr>
<td>Batters hit</td>
<td>1.123 (7.92)***</td>
<td>1.121 (7.38)***</td>
<td>1.141 (2.78)***</td>
</tr>
<tr>
<td>Home runs</td>
<td>0.971 (3.30)***</td>
<td>0.972 (3.03)***</td>
<td>0.962 (1.43)</td>
</tr>
<tr>
<td>Runs ahead and/or behind</td>
<td>1.048 (20.29)***</td>
<td>1.048 (19.08)***</td>
<td>1.052 (6.84)***</td>
</tr>
<tr>
<td>Absolute score</td>
<td>1.01 (2.90)***</td>
<td>1.011 (3.00)***</td>
<td>1.001 (0.10)</td>
</tr>
<tr>
<td>Observations</td>
<td>33676</td>
<td>30284</td>
<td>3392</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-25249.28</td>
<td>-22716.81</td>
<td>-2528.90</td>
</tr>
<tr>
<td>Likelihood ratio $\chi^2$</td>
<td>686.17***</td>
<td>626.73***</td>
<td>66.61***</td>
</tr>
<tr>
<td>Wald $\chi^2$</td>
<td>718.48***</td>
<td>658.55***</td>
<td>71.57***</td>
</tr>
</tbody>
</table>

NOTE: Incidence rate ratios with robust $z$ statistics in parentheses. Constants and year effects not reported.

*p < .10. **p < .05. ***p < .001.
NOTES

1. The pitcher is an agent of the team who shirks his responsibility to his principals in response to imperfect monitoring. Teammates may experience retaliation and perform suboptimally at the plate in trying to avoid expected retaliation. Managers suffer from suboptimal performance of players and are subject to fines, suspensions, and ejections.

2. In a follow-up to the original article, Goff, Shughart, & Tollison (1998) point out that control variables in their analysis proxy for the batter composition effects.

3. The likelihood that the hit batsmen rates of the leagues are equal from 1921-2003 is approximately 1 in 1000.

4. Goff et al. (1998) first suggested the dilution of talent from expansion as a possible reason for the change in the 1990s. The rules of the 1993 expansion draft allowed AL teams to protect more players than NL teams. In addition, the expansion teams were added only to the NL, not the AL. In 3 of the 5 years following the expansion, the NL rate of hit batsmen exceeded the AL rate. In the 1998 expansion, there was no such asymmetry in the expansion draft rules, and each league received an expansion team. Since the second expansion, only once (2000) has the NL hit batter rate exceed the rate of the AL. See Pappas (1997) for a summary of the expansion draft rules.

5. The change, which modified Rule 8.02(d) of the Official Baseball Rules of MLB, was a part of the new On-Field Behavior Policy adopted in December 1993 (National Association of Professional Baseball Leagues, 1996). We would like to thank Jim Porter for identifying the exact timing of this rule change.

6. A similar rule change in Japanese professional baseball led to a narrowing of hit batsmen across its two leagues, one that allows the DH and one that does not. See Kawaura and LaCroix (2002) for details.

7. The Retrosheet game logs have data going back to the earliest days of baseball; however, before the early 1970s the game data is quite sporadic and not complete.

8. To be safe, we also estimated the model using the negative binomial regression. The results were not substantially different from the Poisson estimates.

9. The average seasonal difference in hit batter rates between DH and non-DH games, without respect to league, is approximately 16%.

REFERENCES


**John Charles Bradbury** is an assistant professor of economics at Sewanee: The University of the South. He has a Ph.D. in economics from George Mason University and BA from Wofford College. His research interests include public economics and regulation. Some significant publications include “Legislative Organization and Government Spending: Cross-Country Evidence” (with M. Crain) in *Journal of Public Economics* and “Bicameral Legislatures and Fiscal Policy” (with M. Crain) in *Southern Economic Journal*.

**Doug Drinen** is an assistant professor of mathematics at Sewanee: The University of the South. He has a Ph.D. in math from Arizona State University, and his field is operator algebras. His other current research interests include the mathematics and economics of sports. Some significant publications include “The C*-Algebras of Arbitrary Graphs,” in *Rocky Mountain Journal of Mathematics*, and “Viewing AF-Algebras as Graph Algebras,” in the Proceedings of the American Mathematical Society.