The Designated Hitter Rule in Baseball as a Natural Experiment

Akihiko Kawaura
Dept. of Policy Studies, Doshisha University
Kyoto, Japan

Sumner La Croix¹
Dept. of Economics, University of Hawaii
Honolulu, Hawaii

Working Paper 10-5
April 7, 2010

Abstract

All but two professional baseball leagues have adopted the “designated hitter” (DH) rule, which allows a team’s manager to designate a player to bat at the plate and run the bases in place of another player, usually the team’s pitcher. Unlike the team’s other players, the designated hitter does not take the field to play defense. This paper provides a survey of a large literature investigating the DH rule’s effect on the incentives of pitchers to hit batters and on changes in the number of hit batsmen. We also consider whether the DH rule provides a good example of a natural experiment, as some professional baseball leagues were “treated” with the DH rule and others were not treated.

JEL codes: L83, D82

Keywords: baseball, designated hitter, moral hazard, natural experiment, Japan

¹ We thank Sang-Hyop Lee for helpful comments on the paper. The authors are responsible for all errors and omissions.
The game of baseball has been played with a set of playing rules that have varied little across countries and over time. Nine players play nine innings; each team bats until it makes three outs; three strikes is an out; and so on. There has, however, been one major revision to the rules of play that all but two professional baseball leagues have adopted. Known as the “designated hitter” (DH) rule, it allows a team’s manager to designate a player to bat at the plate and run the bases in place of another player, usually the team’s pitcher. Unlike the team’s other players, the designated hitter does not take the field to play defense. In modern professional baseball, where skill specialization is rigorously pursued, the designated hitter is an offensive specialist, whereas a pitcher is a defensive specialist.

To some fans and regular observers of the game, the DH rule represents “a radical departure from the game’s formerly fundamental principal that every player should both come to bat and play a position in the field” (Zorn and Gill, 2007, p. 192). To many sports economists, the DH rule represents a change in the constraints and opportunity costs faced by players, team managers, and general managers. In particular, there have been numerous studies regarding the DH rule’s effect on the incentives of pitchers to hit batters and on changes in the number of hit batsmen. Section I provides a survey of this literature.

To some sports economists, the change to the DH rule is a good example of a natural experiment, as some professional baseball leagues were “treated” with the DH rule and others were not treated. The coexistence of multiple treatment and comparison groups could allow researchers to identify and estimate the effects of the introduction of the DH rule on fan, team, and player actions. Section II examines whether the introduction of the DH rule is a good example of a natural experiment. Section III concludes with some thoughts about how this research has contributed to our understanding of baseball and, more generally, of economics.

I. DEBATE ON DH RULE, HIT BATSMEN AND MORAL HAZARD

In the United States and Canada, the American League of Major League Baseball

---

2 Describing the debate between pros and cons of the DH rule as a contest of “change” and “tradition”, Zorn and Gill (2007) investigate whether people’s opinion about the DH rule reflects their socio-political backgrounds. Using data from a media monthly poll conducted in September 1997 by CBS News, they estimated a logit regression of approval of the DH rule on respondents’ political orientation, age, sex, and race. Their main finding is that Democrats are 90 percent more likely to support the DH rule than independents and Republicans (p. 193).
(MLB) adopted the DH rule in 1973. In Japan, the Pacific League of the Nippon Professional Baseball (NPB) adopted the rule in 1975. One MLB league—the National League—and one NPB league—the Central League—never adopted the rule. Professional baseball leagues in South Korea, China, and Taiwan all use the DH rule.

Numerous researchers have investigated the DH rule’s effect on the number of hit batsmen in the context of pitchers’ moral hazard using MLB and NPB data. The debate was prompted by Goff, Shughart and Tollison (1997), hereafter GST (1997), who hypothesized that moral hazard by pitchers leads to more hit batsmen in a league with the DH rule. When a pitcher hits an opposing batter, he may face retaliation at his next plate appearance as a batter. Thus he thinks twice even if he is tempted to throw at some player in the opposing team’s batting line-up. The DH rule removes this personal retaliation potential because the pitcher no longer has to bat. GST (1997) argued that the DH rule thus shifts the balance in pitchers’ cost-benefit calculation of hitting the opposing batters. The cost of throwing at batters of their personal choice declines, and they respond by hitting more batters.

Using season data for the American and National Leagues from 1901 to 1990, GST (1997) found that batters in the American League were hit by pitches more frequently than their National League counterparts from the 1973 season when the American League adopted the DH rule. In regressions that explain annual inter-league hit batsmen differences by the use of the DH rule and differences in at-bats, they find that American League batters were hit by pitches about 44-50 times more than their counterparts in the National League after the American League adopted the DH rule. This represents an increase of 10-15 percent in the number of hit batsmen per season.

Levitt (1998) casts doubts on the moral hazard hypothesis. Using season data for the National and American Leagues for the 1993-1996 seasons, he found that pitchers were rarely hit by pitches in the National League, which suggests that they are not necessarily a target of retaliation. Data covering 60 National League pitchers during the 1993-1996 seasons does not support the hypothesis that pitchers who frequently hit opposing batters were hit more often by opposing pitchers. Trandel, White and Klein (1998), hereafter TWK (1998), also questioned the moral hazard story by arguing that a pitcher merely acts as an agent of his team’s manager and that retaliation is rarely directed at pitchers themselves. Instead, retaliating against a team’s big slugger provides greater benefits.

---

3 The American League owners voted 8-4 on January 11, 1973 to experiment with a three-year trial of the DH rule.
Levitt (1998) and TWK (1998) offered an alternative explanation for the surge in hit batters that appears to have been prompted by the DH rule. They argue that the DH rule brings about a stronger batting line-up, as weak-hitting pitchers are replaced by designated hitters who are usually big sluggers. Since the opportunity cost to a team of hitting a batter is lower when the batter is a good hitter, the presence of the DH in the lineup will lead to more hit batters. TWK (1998) amended GST’s (1997) regression specification to allow for hit batters to be proportional to at-bats and extended their MLB sample by seven seasons (1991 to 1997). In the new regression specification, the DH effect was no longer statistically significant. In fact, TWK (1998) found that the DH effect was not detected with GST (1997)’s original specification in their extended sample.

In response to these criticisms, Goff, Shughart and Tollison (1998), hereafter GST (1998), acknowledged that the DH effect is not detected when the sample period is extended through the 1997 season. They hypothesized that the entry of two teams into the National League in 1993 may have diluted National League pitching and led to more hit batters in the league, thereby diminishing the size and statistical significance of the DH effects in the American League. They also speculated that the players’ strike in 1994-1995 may have affected regression results with the extended sample.

Trandel (2004) focused on the incentives to retaliate and investigated whether retaliation ever exists, be it against offending pitchers themselves or against their teammates. He argued that if a team retaliates when its own batter has been hit, there should be a positive correlation between the number of opposing batters hit by the team’s pitchers and the number of hit batters on the team. He acknowledges that retaliation in a broad sense could take the form of delayed revenge against the plunking pitcher or his teammates in future games between the teams. To allow for this possibility, Trandel (2004) examined season data on hit batters by MLB team from the 1960 to 2002 seasons. He calculated the correlation between the deviation from the league average of a team’s own batters hit by pitches and the deviation from the league average of a team’s opposing batters hit by pitches. The correlation coefficient (-0.018) is negative, very small, and statistically insignificant. Since opposing batters hit by the team’s pitchers are not positively correlated with team batters hit by opposing pitchers, Trandel concludes that there is no evidence for retaliation.

The studies reviewed so far mostly used aggregate season data by league (or by team) to investigate the impact of the DH rule in terms of incidence of retaliation and hit batters. Bradbury and Drinen (2006) changed the terms of the debate by exploiting MLB
game-level statistics. Their sample encompasses 131,342 MLB games played over the course of the 1973-2003 seasons. Their regression analysis of individual game statistics provides strong evidence that retaliation does exist: Each opposing batter that a team’s pitcher hits raises the team’s own hit batters by 10 to 15 percent in the same game. Controlling for various factors that include team-specific variables such as batter quality, pitcher quality, and game-specific score difference variables, an American League game played with the DH is still associated with 5-12 percent more hit batters than a National League game played without it. Bradbury and Drinen’s results provide strong support for the moral hazard hypothesis grounded in the retaliation threat.

As Bradbury and Drinen’s sample includes data from interleague games that began in 1997, they were able to use these games to remove the potential effect of league-specific factors. In interleague games, use of the DH rule is not restricted to teams from the American League. When inter-leagues games are played in the American League team’s stadium, then both teams are permitted to use a DH; when games are played in the National League team’s stadium, neither team is permitted to use the DH. Estimation of the impact of the DH rule on hit batters with just the interleague sample yields a positive DH effect on hit batters, thus reinforcing the moral hazard story.

Bradbury and Drinen (2007) brings the analysis of this question to new heights by using a sample composed of micro-level play-by-play MLB data sets with 641,640 observations for the 1989-1992 seasons and 584,886 observations from the 1969 and the 1972-1974 seasons. Using probit and logit regression models, they regressed a batter being hit by a pitch during a particular time at bat on the DH dummy and various variables representing game situations. Regression results were similar for both data sets, with the DH rule and higher batter quality associated with a higher probability of the batter being hit. Their most striking finding is that “[T]hough pitchers are hit less frequently than other players on average, pitchers are more likely to be hit after plunking an opposing player” (p. 137).4

Kawaura and La Croix (2007) investigates the effect of the DH rule with data from Japan’s professional baseball leagues. The adoption of the DH rule among Japanese

4 Stephenson (2004) examined whether starting pitchers and relief pitchers hit batters at different rates after switching leagues. His analysis of data for 206 pitchers who switched leagues did not find statistically significant differences in hit batters rates for either group of pitchers. He concluded that his results do not support the moral hazard hypothesis.
professional leagues parallels the MLB split and should provide independent tests of hypotheses relating to the DH rule and additional insights on the DH debate. The sample consists of data from the 1958-2004 seasons by team: Six teams from the Pacific League which began to use the DH in 1975, and six from the Central League which has never used the DH. An absence of entry and exit of teams in Japan during the sample period provides a more stable environment for the empirical analysis than in the United States.5

Kawaura and La Croix replicates analyses of GST (1997) and TWK (1998) using aggregate data from the two Japanese leagues and found that the DH rule was associated with more hit batsmen until a strict rule was imposed against “dangerous pitching” in 1989.6 We accounted for the batter quality with team data, and still found that the DH rule led to more hit batsmen, thus lending support to the moral hazard story. One innovation of our analysis is the introduction of team defensive strategy in the regression analysis. A manager of a team with poorly performing pitchers may instruct his players to engage in aggressive pitching to compensate for their lack of talent. This strategy is easier to implement when pitchers do not have to fear personal retaliation. Our regression analysis found that the impact of team pitching talent on the number of hit batsmen became larger when the DH rule was introduced. Even after controlling for a team’s defensive strategy, we still found that the DH rule was associated with a small number of additional hit batsmen. We conclude by acknowledging that it is difficult to distinguish econometrically whether increases in hit batsmen were due to moral hazard by pitchers or to changes in team defensive strategy.

Kawaura (2010) focuses on individual pitchers, and questions whether their calculation of the costs and benefits associated with hitting a batter shifted in the direction discussed in GST (1997). He argues that if a pitcher begins to throw at opposing batters as soon as the DH rule is in place, it could embarrass his fans. To the extent that team owners are concerned with ticket revenues, loss of fans would in turn lower the demand for his service among them. There could also be fewer off-field commercial opportunities such as public appearances and product endorsements. When these costs of hitting batters are large

---

5 Some teams did switch cities during the sample period.

6 The 1989 rule change in Japan gave umpires the authority to remove the pitcher (and his manager) when they judged that the pitcher engaged in dangerous pitching. Kawaura and LaCroix (2007) assigned separate dummies for two DH periods with and without the “dangerous balls” rule. A similar rule was introduced in MLB in 1994; it provided for a “double-warning” to both teams against intentional hit batsmen and provided authority to remove a pitcher intentionally throwing at batters. The DH rule literature based on the MLB data has, however, not incorporated its effect in the empirical analyses of determinants of hit batsmen.
enough, a pitcher may resist the temptation to plunk a batter even if he need not be afraid of retaliation.

As this reputation effect is assumed to be of greater relevance for established pitchers, Kawaura investigates behavioral changes of six Pacific League pitchers who played consecutively during the 1969-1980 seasons. Thus, each of these pitchers played six seasons without the DH rule and six seasons with the DH rule. Panel analysis of the number of batsmen hit by these pitchers did not produce results consistent with the moral hazard hypothesis, as the number of batters hit by these pitchers did not increase even after they no longer had to appear at the plate in the offensive capacity. Shortening the sample period to ten years to increase the number of eligible pitchers to eleven did not alter the regression results.

Kawaura argues that the absence of behavioral changes among veteran players who experienced the adoption of the DH rule midway through their career is compatible with the DH effect on hit batsmen previously identified in the DH rule debate. This is because pitchers who began to play when the DH rule was already in effect did not have a reputation to protect and thus could hit batters more frequently compared with pitchers who played in the pre-DH period. A changeover of pitcher composition from the latter to the former could have raised the aggregate number of hit batsmen per team even when veteran pitchers did not alter their behavior.

II. IS THE DH RULE A GOOD NATURAL EXPERIMENT?

The studies that attempt to identify the DH rule’s effect on hit batsmen typically compare data generated in the league that allows the use of the DH with those from the non-DH league. One could argue that this comparison constitutes a natural experiment, as the non-DH league serves as the control group, while the DH league constitutes the “treated” sample. This section evaluates the validity of this claim.

We begin by reviewing criteria that should be met for the data sample to be considered the result of a natural experiment. Meyer (1995) and Rosenzweig and Wolpin (2000) provide two excellent surveys of the natural experiment framework as applied by economists. Meyer defines a “good natural experiment” as one “in which there is a transparent exogenous source of variation in the explanatory variables that determine the treatment assignment” (p. 151). Typically, there are available data both before and after the treatment and there is an
untreated comparison group that is exposed to the same types of influences as the treated group. Studies tend to be more reliable when there is less attrition (or addition) of observations in either the treatment or control groups. Meyer (p. 157) argued that the experimental design can be strengthened when there are multiple comparison groups.

Economists have studied a wide range of topics with the natural experiment framework, including the effect of job training programs on earnings, the effect of class size on standardized test outcomes, and the effect of exogenous changes in fertility on the labor supply of married women. Angrist’s (1990) study of the effect of veteran status on individual earnings is a good example of a natural experiment. Estimating how veteran status affects earnings is complicated because individuals do not randomly select military service. Individuals with poor health (and therefore potentially lower incomes) are screened out by the military, while those who enlist in the military may have different labor market opportunities than those who do not serve. Angrist (1990) overcame the second selection problem by examining the labor market outcomes of men who were randomly drafted into the army during World War II and the Vietnam War. Induction into the Army during the Vietnam War was based on a draft lottery, while induction during World War II was based on a man’s date of birth. Thus, he argues that regression estimates of the effect of veteran status on earnings using this sample are more likely to provide unbiased estimates of the effects of veteran status on earnings.

How well does the natural experiment framework fit the adoption of the DH rule by one of two leagues in North America and Japan? Several conditions are clearly met. First, the nature of competition between teams in both the treated and the control leagues are virtually identical but for the DH rule. Second, there are multiple years of data (by season) on player, team, and fan behavior prior to and after the adoption of the DH rule. Third, there is no attrition by teams in the treated and control groups, although in North America one

\[ y'_{it} = \alpha + \alpha_t d_t + \alpha_j d_j + \beta d_{jt} + \delta j + \epsilon_{it} \]

where the outcome \( y \) is now also indexed by \( j \) for the group, \( j = 0, 1 \), and \( d_t = 1 \) if \( t=1 \) and 0 otherwise, \( d_j = 1 \) if \( j=1 \) and 0 otherwise, and \( d_{jt} = 1 \) if \( t=1 \) and \( j=1 \) and 0 otherwise. \( z_{it} \) is a vector of the observations’ characteristics controlled for in the regression. \( d_j \) is a dummy variable for being in the experimental group after it receives the treatment, and \( \beta \) is the true causal effect of the treatment on the outcome for this group. The key identifying assumption is that \( \beta \) would be 0 in the absence of the treatment, or \( E[\epsilon_{it} | d_{jt}] = 0 \). See Meyer (1995) for a full discussion.

While the army randomly chose draftees in both wars, it continued to screen out some men with poor health. Other individuals obtained draft deferments because they were enrolled at institutions of higher education.

---

7 The standard regression is \( y'_{it} = \alpha + \alpha_t d_t + \alpha_j d_j + \beta d_{jt} + z_{it} \delta + \epsilon_{it} \) where the outcome \( y \) is now also indexed by \( j \) for the group, \( j = 0, 1 \), and \( d_t = 1 \) if \( t=1 \) and 0 otherwise, \( d_j = 1 \) if \( j=1 \) and 0 otherwise, and \( d_{jt} = 1 \) if \( t=1 \) and \( j=1 \) and 0 otherwise. \( z_{it} \) is a vector of the observations’ characteristics controlled for in the regression. \( d_j \) is a dummy variable for being in the experimental group after it receives the treatment, and \( \beta \) is the true causal effect of the treatment on the outcome for this group. The key identifying assumption is that \( \beta \) would be 0 in the absence of the treatment, or \( E[\epsilon_{it} | d_{jt}] = 0 \). See Meyer (1995) for a full discussion.

8 While the army randomly chose draftees in both wars, it continued to screen out some men with poor health. Other individuals obtained draft deferments because they were enrolled at institutions of higher education.
of the two leagues—the National League—adds teams during the period after the American League adopted the DH rule. Finally, there are essentially two parallel experiments occurring in two very distant regions of the world; this provides a check on singular results that might be generated by the DH rule in a particular country.

Some of the assumptions undergirding the natural experiment framework are, however, not supported in either North America or Japan. Rosenzweig and Wolpin (2000) argued that the major problem with most empirical studies that claim to be natural experiments is that “the assumption of randomness [of the treatment] is not credible” (p. 828). This assumption is also violated in studies of the effect of the DH rule that assume that adoption of the DH rule is a random treatment. We argue that the DH “treatment” was not randomly assigned to particular leagues, but was instead selected by both the American League and the Pacific League. Attendance in both leagues had been stagnating relative to their sister professional baseball leagues, and baseball team owners and general managers had been considering changes in the game to increase interest in baseball. A comparison of attendance in the American and National Leagues before and after the introduction of the DH rule in 1973 shows that annual attendance per team in the American League increased rapidly after the introduction of the DH rule and that the two attendance series closely co-vary from 1978 (Figure 1). A parallel comparison of attendance in the Central and Pacific League before and after the introduction of the DH rule in 1975 shows a less clear relationship between adoption of the DH rule and league attendance (Figure 2). There is, however, a change from no growth in attendance in the Pacific League over the decade prior to the DH rule adoption to small positive growth over the next decade (1975-1984), followed by high growth over the next six years (1985-1990). While the evidence is stronger for North American baseball, the introduction of the DH rule is associated with changes in the growth rate of attendance in North American and Japanese professional leagues.

Rosenzweig and Wolpin (p. 828) identify five naturally occurring phenomena that “are plausibly random with respect to at least two of the major sources of heterogeneity in human populations: tastes and abilities”. They are “twin births, human cloning (monozygotic twins), birth date, gender, and weather events.” Rosenzweig and Wolpin (p. 872) concluded their survey of econometric studies using natural instruments on a clearly cautionary note: “It is evident that natural events used as instruments do not provide estimates that can be unambiguously interpreted, although the range of possible alternative interpretations may have been reduced.”

In 1969, both the National and American Leagues lowered the height of the mound from 15 to 10 inches.

Other events also affected attendance in the American and Pacific Leagues in the late 1960s and early 1970s. La Croix and Kawaura (1995) found that the introduction in 1966 of a new system for allocating young players to teams—the player draft—substantially increased a number of measures of
Another problem with viewing the change to the DH rule as a natural experiment is that the composition of players in each league changed after the introduction of the DH rule. Teams in the league with the DH had incentives to trade good-hitting pitchers to teams in the other league that would place a higher value on their hitting skills. Tollison and Vasilescu (2007) developed and tested this hypothesis using data on the offensive and defensive performance of 6,052 MLB pitchers for the 1960-1985 seasons. A principal finding is that National League teams were more likely to trade poor-hitting pitchers to the American League in the two seasons (1972 and 1973) bracketing the introduction of the DH rule in April 1973. This association is not observed before 1972 or after 1973. One implication of Tollison and Vasilescu’s results is that the existence of these trades makes it more difficult to view the introduction of the DH as a natural experiment, as the skills of the pitchers in the National League changed in anticipation of the rule change.

Kawaura and La Croix (2010) examine whether there was a similar surge in inter-league pitcher trade activities in Japan’s two professional leagues in the years leading up to and including the adoption of the DH rule by the Pacific League in 1975. For example, trades of pitchers from teams in the Central League to teams in the Pacific League averaged 2.93 over the 1958-1971 period; they increased substantially, to 9.25, over the 1972-1975 period. Trades of pitchers from teams in the Pacific League to teams in the Central League also increased, from an average of 2.93 in the 1958-1971 period to 9.25 over the 1972-1975 period.

Following Tollison and Vasilescu’s hypothesis, were the pitchers traded to the Central League better hitters than the pitchers traded to the Pacific League over the 1972-1975 period? Kawaura and La Croix (2010) find that the average on-base percentage and the average slugging average (OPS) for pitchers traded to the Central League was higher than for pitchers traded to the Pacific League. A \( t \)-test that the two league averages are equal is rejected at the 10 percent level of statistical significance.

There is also some evidence that National League attendance would not have reacted in the same way to the introduction of the DH rule as American League attendance.\(^\text{13}\)

---

\(^{12}\) Since American League pitchers had no offensive opportunities after 1973, their batting ability cannot be observed and they had to be dropped from the sample after this year.

\(^{13}\) In other words, the effect of treatment on the control group would not have been the same as the
Domazlicky and Kerr (1990) estimated a regression model for MLB team attendance using data from the 1969-1980 seasons for 21 U.S.-based MLB teams that did not change their location during the sample period. The estimated coefficient on number of runs scored is positive and statistically significant in the American League sub-sample, while it is statistically insignificant in the National League sub-sample. Since the DH rule would be expected to produce more runs per game, Domazlicky and Kerr’s results provide some insights into the National League’s decision to reject the DH rule and the American League’s decision to adopt it.

Finally, we take note of the addition of new teams to the National League in 1990 while the DH rule was in effect in the American League. Since the presence of new teams reduced the quality of all team rosters for an undetermined period of time, this effect may have contributed to estimates of differences in behavior due to the presence of the DH rule. We note that the two Japan leagues did not add or drop teams between 1958 and 2004, although some teams (as in North American baseball) moved to different cities.14

III. CONCLUSIONS

The DH rule is unique among baseball rules because professional baseball leagues in the United States and Japan have not uniformly adopted it. In the context of a natural experiment framework, researchers have engaged in a long-running debate regarding the relationship between the DH rule and increases in hit batters. The issue is whether the moral hazard of pitchers contributed to increases in hit batters after the DH rule was enacted. Empirical analyses using aggregate season data have come to different results, while studies based on disaggregated game data and play-by-play data clearly revealed pitcher moral hazard.

On a closer look, however, the asymmetry in the status of the DH rule in two countries was not sufficient to produce a true natural experiment, as its adoption by the American and Pacific Leagues was endogenous, an attempt by both leagues to increase fan attendance and interest in their games. While the presence of one league without the DH rule would have provided a natural experiment, the endogenous adoption of the DH rule in both leagues did not create such an environment.

14 Two teams in the Pacific League (Bluewaves and Buffaloes) merged after the 2004 season. There was an entry (Golden Eagles) immediately following this merger, which maintained the number of teams at six in the following seasons.
rule appears to offer a control group for regression analysis, studies of both U.S. and Japanese professional leagues show that inter-league player trades were conducted to optimize team rosters at the time of the adoption of the DH rule by one league in each country.

In sum, the DH rule in the North American and Japanese professional baseball leagues continues to offer opportunities to researchers to study a rule’s effect on various aspects of the sport. For example, the rule allows a partial division of offensive and defensive strategies in baseball, which could change the required talents of a team manager depending on the status of the DH rule in the league. Another potential issue is the recruitment of rookie pitchers to professional teams. Because the DH rule diminishes the importance of hitting talents of pitchers, it could have consequences on the scouting decisions of professional teams as well as on the behavior of amateur players.
References


Figure 1: Attendance in National and American Leagues: 1969-2000
Figure 2: Attendance in Central and Pacific Leagues: 1958-1993