Opportunistic Shirking and MLB Umpires

Michael Lopez, Skidmore College
Brian Mills, University of Florida

Finance/Economics - Economics (Professional Sport)
20-minute oral presentation (including questions)
Abstract 2018-238

Relevance and Literature Review
Considerable attention in sports economics and personnel economics has been devoted to the problem of shirking by employees and the role played by wage structures, incentive schemes, or direct monitoring in enhancing effort (Maxcy & Fort, 2002; Krautmann, 1990; Marburger, 2003; Krautmann & Donley, 2009; Depken, 2000; Berri & Krautmann, 2006; Krautmann & Solow, 2009; Lazear, 2000). As direct monitoring can be costly, various firms turn to the use of wage structures and incentives to encourage their respective employees to perform at their highest ability. Alternatively, employees may also choose to participate in behavior that can signal high effort, such as agreeing to work unpaid overtime. This may be particularly salient among workers seeking promotion or those on temporary contracts seeking permanent employment (Engellandt & Riphahn, 2005; Anger, 2008). The relevance of these inquiries within the management of sports and athletes is clear: teams and leagues want players and officials to offer maximum effort to ensure continued fan interest through providing the highest level of elite athletic competition.

This paper therefore tests the occurrence of shirking behavior in the context of unpaid overtime, where a natural experiment occurs that allows workers to behave in ways that minimize expected overtime work. We set our study in the context of professional umpiring in Major League Baseball (MLB), where employees are salaried and there exists a high firm cost of dismissal. We identify unpaid overtime work using extra innings, which provide a salient break in regular and overtime work, but for which there is no additional compensation paid.

The primary mechanism by which umpires impact game outcomes comes via ball and strike calls. Since 2008, each of these ball-strike decisions has been monitored by a tracking system. Although related research suggests that monitoring mitigates umpire bias (Parsons et al., 2011) there is evidence that biases still exist in the face of this scrutiny (Mills, 2014; Kim & King, 2014, Chen, Moskowitz, & Shue, 2016). There is therefore the possibility of opportunistic shirking using ball-strike decisions, particularly given that extra innings (unpaid overtime) likely represents a limited fraction of an umpires' overall evaluation. If umpires can manipulate the expected length of a game through ball-strike calling behavior, they may choose to do so if the effect that these choices have on their overall employment status or pay is relatively small.

Data and Methods
Extra innings occur only when the game is tied after nine full innings, and allow both the home and away teams an attempt to score additional runs until one team has more runs than the other at the end of a full inning. Therefore, unpaid overtime for an umpire is indefinite and unknown when each inning begins. However, umpires have additional information about the game's expected game length in the bottom half of an inning. If the away team scores and takes the lead, the game ends if the home team fails to score. If play is tied during the bottom of an inning, but the home team has runners on base, the likelihood of the home team scoring to end the game is increased substantially. In each case, the umpire is given the opportunity to influence expected game length through ball-strike calls: increasing strike calls when the away team is leading, or decreasing strike calls in tie games with runners on.

We use three game states to analyze pitches taken during the bottom of extra innings: the away team leading, a tied game with no runners on base, and a tied game with at least one runner on base. We first show that strike rates are significantly different in each game state -- highest when the away team leads and lowest when the game is tied but the home team has at least one runner on base. Next, we use semi-parametric modeling to estimate changes to the probability that a pitch is called a strike, conditional on its location, across game states. We find that for pitches on the border of the strike zone, umpires are up to 16 percentage points more likely to call a strike when the away team
is leading relative to when the home team is tied but threatening to score. Both findings are consistent with umpires
reducing time in the current overtime and the potential for future overtime.

Each MLB pitch from the 2008-2016 regular seasons was collected from Baseball Savant using the baseballR
package in the statistical programming software, R. These data contain pitch and game-level characteristics including
date, inning, inning half, batter handedness (left or right), base state (runners on 1st, 2nd, and/or 3rd base), pitch
location (horizontal and vertical coordinates), and pitch result (ball in play, called ball, called strike). Our analysis
focuses on pitches thrown in the bottom of extra innings. We merge this data with information on umpire home
plate assignments gleaned from Retrosheet. In total, our data include 32,785 pitches thrown in extra innings by MLB
pitchers. We estimate a generalize additive model (GAM) to control for pitch location, and compare the odds of a
strike call conditional on both this pitch location and the game-state as characterized above.

Results and Discussion
Consistent with opportunistic shirking, umpires call more strikes (odds ratio 1.15, 95\% interval 1.09 - 1.22) when
the home team is about to lose, and fewer strikes (odds ratio 0.84, 95\% interval 0.79 - 0.89) when they are about to
win. We note that this initial odds summary does not take into account the location of individual pitches, which may
differ systematically with game-state. Therefore, we use our GAM to account for locational disparities.

Because smoothing parameters in GAMs are difficult to interpret, our presentation will display absolute differences
in called strike percentages by game state and stance, comparing estimates from our GAM models for each game
state. Colors and color darkness will reflect the direction and magnitude of the differences in strike likelihood,
respectively. The results of these estimations reveal a similar pattern to our more simplistic odds calculation.

Across much of the fringe area of the strike zone, both loss imminent (higher chances, up to 12.0 percentage points
more) and win imminent (lower chances, up to 10.1 percentage points less) game states differ from neutral game
ones, with differences for win versus loss imminent states as large as 16.4 percentage points. Each of these effects is
consistent with behavior that would be expected to shorten games and reduce the likelihood of additional extra
innings (overtime). Pitches in the middle of the strike zone, as well as outside the strike zone, show no obvious called
strike differences, making it clear that umpires exert their discretion on pitches that are more easily disguised as
correct calls. Both the consistency of the results across fringe areas of the strike zone, as well as the magnitude of
strike call differences identified with the same type of data are notable. From the standpoint of statistical significance,
the GAM estimation appears to offer a substantially better fit relative to GAMs that do not include game state
indicators, as judged using likelihood ratio tests and the Akaike Information Criterion.

Contribution to the Body of Knowledge
We show that umpires tend to make ball-strike calls consistent with intentions to end games earlier and reduce time
worked. These results are important for leagues with interest in maintaining the integrity of competition that attracts
fans to the stadium and to view broadcasts on television. As information related to ball-strike call accuracy becomes
more salient to fans – for example, through real-time on-screen pitch tracking visuals – MLB may be well-served to
ensure that its officials are acting in a way consistent with the stated rules.

Alternatively, if fans are less likely to remain tuned into extremely long games, MLB may have an incentive to
encourage umpires to shorten these types of games when this opportunity arises. Clearly, there are important
considerations to be made at the league level to maximize fan interest and, in turn, league revenues. While we do not
provide a recommendation for the optimal choice of bias levels for revenue maximization, the results here clearly
indicate that umpires act in biased ways in certain situations that favor one team over another.