# The Dynamics of Inattention in the (Baseball) Field

## Online Appendix*

James Archsmith[†]
University of Maryland, College Park

Anthony Heyes[‡]
University of Birmingham

Matthew Neidell[§]
Columbia University

Bhaven Sampat[¶]
Arizona State University

Draft Date: January 21, 2025

**Preliminary and incomplete**

Please do not circulate or cite without permission of the Author.

Please review and cite the most recent version of this paper available at
https://econjim.com/WP2101

*We are grateful to Stephen Coussens, Jonathan Guryan, Sandeep Kapur, Adam Leive, Devin Pope, Stephen Salant, four reviewers from this journal and seminar participants at the University of Virginia, University of Massachusetts and University of Connecticut for helpful comments and suggestions.

†Department of Agricultural and Resource Economics, University of Maryland, College Park, MD 20742 Email: archsmit@umd.edu, URL: https://econjim.com, ORCHID: 0000-0002-6052-0302

‡Department of Economics, University of Birmingham, Birmingham, UK Email: a.g.heyes@bham.ac.uk ORCHID: 0000-0003-1847-9374 Heyes acknowledges financial support from the Canada Research Chair program.

§Department of Health Policy and Management Columbia University, New York, NY 10032, NBER and IZA Email: mn2191@columbia.edu

¶School for the Future of Innovation in Society, Arizona State University, 1800 I Street NW, Washington DC 20016 and NBER. Email: bhaven.sampat@asu.edu ORCHID: 0000-0001-7387-6706

# Online Appendix

## A Robustness

### A.1 Results Excluding the Ninth Inning

Since we find a significantly different impact of the 9th inning on umpire decision accuracy, and the 9th inning may have particularly different leverage, we seek to explore whether this inning has an outsized impact on our results. As test of robustness, we re-estimate our primary model, limiting our sample to innings two through eight. Results are shown in the tables below. Estimated parameters are not substantially different from models where we include the ninth inning, suggesting any impact from the 9th inning does not affect our main conclusions.

[Table A.1 About Here]

### A.2 Results Using Actual Instead of Simulated Leverage

Our measure of the leverage at each pitch requires that we compute two probabilities in each game state: the probability a given team wins in the event of a called ball and the probability they win if there is a called strike. Although we use simulated data, it is possible to empirically calculate these probabilities using observed outcomes in MLB games. However, while we have a wealth of data on which to base these estimates (over three million pitches in over 26,500 games), the space of possible states is also large, and using actual game data to compute leverage could lead to substantial measurement error.[1] To address this concern, the leverage measure used as the basis of our primary specifications is derived from simulations of 5 million MLB games.

A test of the robustness of using this simulated leverage metric would be to re-estimate our primary specifications using leverage computed using a leverage measure based on game data-only (GDO), with the understanding that it is poorly measured. Assuming classical measurement error, we would expect the parameter on leverage to be biased toward zero in these estimates.[2]

As a prerequisite, we first compute the degree of attenuation bias that can be expected given the measurement error in the game data-only leverage measure. Assuming the simulated leverage measure is the "true" measure of leverage on every pitch, attenuation bias is a function of the variance of the true measure divided by the sum of the variance of the true measure and the variance of the error in the noisy

[1] Accounting for all possible combinations of balls, strikes, outs, baserunner positions, inning and inning part, and score differences between a 10-run advantage and a 10-run disadvantage, there are 108,864 possible states. Some states occur very frequently, e.g., every game starts in an identical state, and some states are not observed at all. A given state is observed, on average, around 30 times over the course of our data. If the probability of a team winning were 0.50, the estimated probability based on 30 observations would have a standard error of approximately 0.09. The standard error in leverage estimated this way would be even larger since it is the difference of two such probabilities measured with error. Computing the standard error of leverage requires knowledge of the covariance in the two estimated win probabilities. Applying the Cauchy-Schwarz inequality, we can bound the standard error of a leverage measure based on two outcomes observed thirty times each to [0.1286,0.1296].

[2] If measurement error in the game data-only measure arises only from sampling variation in the estimated win probability in each state, then the measurement error is orthogonal to any other unobserved variables in our regression and meets the criteria of classical measurement error.

measure. Using estimates of these variances from our observed data, we can then compute the expected ratio of parameters from our preferred specification to those from a specification using the GDO leverage measure.

Measurement error in the game data-only measure is driven by sampling variance. Therefore, values of the leverage metric based on game states that are observed more often should be more precise. A natural approach to reducing measurement error would be to limit the sample to states that are observed more frequently, with lower sampling variance. Table A.2 recomputes the variance in the true leverage measure ("signal") and the sampling variance ("noise"), limiting the sample to observations where the game data-only measure is computed using at least 1, 100, 250, 750, 1000, and 2500 observations. Then, using these variances, we compute the ratio between the "true" parameter and the expected value of a parameter when the corresponding independent variable is measured with error.

[Table A.2 About Here]

The results of this table demonstrate the large impact measurement error might have on the estimated model parameters. Using the full sample, the true effect of leverage would be over 8 times value of the parameter one would expect to observe given the magnitude of measurement error in the GDO leverage measure. This bias decreases steadily as we limit the sample to states observed more frequently, but is still over 3 times the value when considering states occurring over 250 times in our sample.[3]

We follow by estimating the impact of leverage on the umpire making a correct call using the GDO leverage measure. That is, for each pitch in each of these games, we define the state as the score differential[4], current inning[5], inning part, number and position of baserunners, number of outs, number of balls, and number of strikes. The estimated probability of the home team winning conditional on that state is the proportion of games where that situation arose to games where that state arose and the home team won.[6] We then compute the leverage in some state $A_t$ as the difference in win probability for the situation $A_t$ incremented by a ball and the state $A_t$ incremented by a strike. This method of computing leverage requires minimal assumptions, only that events in a baseball game follow a Markov process with a state defined by the game state variables. However, despite our large dataset consisting of over 26,000 individual games, the large state space leaves some relevant states unobserved or so infrequently observed that the win probabilities for these states are poorly estimated.[7]

[3] As the minimum number of games increases over 750, the bias factor increases. This results from the fact that while increasing the number of games threshold reduces sampling variance (noise), it also reduces variance in the true leverage measure (signal) as the set of game situations in the sample decreases.

[4] We limit to cases where the score difference between teams is 10 or less.

[5] MLB games that are tied after nine innings continue one inning at a time until the tie is broken at the end of the inning. Consistent with our assumption that states evolve as a Markov process, we treat any inning after the 9th inning as the 9th for the purposes of computing the state.

[6] If a given situation occurs multiple times in a game – which frequently occurs when a batter hits a foul ball with two strikes – it is only counted once for the purposes of this calculation.

[7] The state space consists of 21 possible run differences, nine innings, two inning parts (called half-innings), three outs, eight possible arrangements of runners on the bases, three strike states, and four ball states. This is a total of 108,864 possible states. A typical game will pass through around 300 unique states. Given some states are more likely to occur than others (e.g., the state in the top of the first inning, tied game, zero base runners, balls, and strikes occurs in every game) there is incomplete coverage of the state space.

To further investigate the role of measurement error, we again limit the sample to cases where the GDO measure is computed using data from at least 1, 100, 250, 750, 1000, and 2500 unique occurrences in MLB games. For comparison, we also estimate the leverage effect using the same sample of observations and our leverage measure derived from simulated games. Finally, we compute the bias factor dividing the coefficient on the simulated measure to the coefficient on the GDO measure. The results are shown in Table A.3.

[Table A.3 About Here]

These estimates demonstrate two key advantages to using the simulated measure. First, the discrepancy between results using the simulated and GDO measure are broadly consistent with the magnitudes estimated in Table A.2, declining to approximately 300% when limiting to cases where the GDO measure is based on at least 2500 observations. Second, limiting the sample to cases where the GDO measure is based on more observations increases the magnitude of the estimated coefficient in the simulated leverage regressions. States with few or many underlying observations on which to base the calculation of the GDO leverage measure are not randomly assigned and limiting the sample in this way can bias the estimated coefficients. Using the simulated leverage measure avoids both issues. While this table reveals discrepancies between estimates, the qualitative results still hold.

### A.3 Results Computing Past Leverage Over Pitches Requiring a Decision

We compute our measure of past leverage as the sum of leverage for all previous pitches during the current half inning. Alternatively, one may consider computing leverage over only pitches where the umpire was required to make a decision. We repeat estimates using this alternative measure in Table A.4. Column 1 shows our primary estimates. Column 2 replaces the past leverage measure with an alternative computed over only pitches where an umpire made a decision. Column 3 includes this alternative past leverage measure and the past leverage over all pitches where the umpire did not make a decision.

[Table A.4 About Here]

It is important to note this alternative past leverage measure is mechanically smaller in magnitude than when computed over all pitches since there are fewer observations over which to compute the sum. Taking this change of scale into account, this alternative past leverage measure has no meaningful impact on our results. Interestingly, Column 3 shows only leverage on pitches where umpires are required to make a decision has a meaningful impact on their accuracy. We, however, prefer the past leverage metric computed over all pitches as it is computed similarly to the expectation of future leverage, making the parameter estimates directly comparable.

### A.4 Results By Game Status

One might be concerned that umpires make mistakes that favor the losing team, and this may be correlated with the leverage of a given series of pitches. In Table A.5 below we repeat our primary specification across three subsamples of the data: pitches where the batting team is leading in the score, behind, or tied. Leverage effects are weaker when the batting team is losing, but have identical sign and significance compared to estimates from the full sample.

[Table A.5 About Here]

### A.5 Results in Extra Innings

A typical baseball game consists of nine innings. If a game is tied after the ninth inning, the game proceeds to "extra innings" where the game is extended by an additional inning. If the tie is broken at the end of that additional inning, the game ends, otherwise the process repeats until the tie is broken.[8] Our primary specification is estimated using only innings 2 – 9. As a test of robustness, we repeat our primary specification using only data from these extra inning situations. The results are shown in Table A.6. In both sign and significance the parameters are similar to our primary specification.

### A.6 Impacts of Large Crowds

A recent paper by Dean (2024) provides experimental evidence that noise can reduce cognitive function, raising the possibility that umpire performance is negatively affected by crowd noise. The inclusion of game fixed effects should mean that any factors invariant within a game, for example noise levels that carry throughout a game, would not impact our conclusions. However, a reader might be concerned that the noise level from a large crowd might vary in a way that is correlated with leverage, confounding inference. In Appendix Table A.7 we re-estimate our main specification but adding regressors that interact attendance at a game (as a proxy for noise) with our leverage measures. The estimated coefficients on these interactions are small, and statistically insignificant at conventional levels. Further their inclusion does not disturb other primary estimates meaningfully, so we do not think they are an important source of confounding.

## B Data

### B.1 Calculation of the simulated leverage metric

Our leverage metric is intended to capture the influence a particular ball/strike decision by an umpire will have on the eventual outcome of the game. Following Tango (2006) this metric is the potential change in the probability the home team wins (Win Expectancy or WE) between a future where the umpire calls a pitch a strike versus calling it a ball. Calculating this metric requires knowing the WE in these two potential futures.

Our first step in computing this metric is to define a game state -- a set of current game attributes which define identical situations in games for the purpose of computing the WE. Tango (2006) defines

[8] In an effort to reduce the potentially long duration of tie games, MLB adjusted the extra inning rules temporarily during the COVID season of 2020 and starting permanently in 2023, placing a runner on second base at the start of each extra inning. These rule changes occurred after the period used in our sample.

define game states as the combination of inning (1-9), inning part (top or bottom), number of outs (0-2), baserunner positions (occupied or not for 1st, 2nd, and 3rd bases), difference between the scores of the home and away teams (-4 to 4). We extend Tango's game state in two ways. First, we are concerned with how WE will evolve after individual ball/strike calls by the umpire so we add the current ball (0-3) and strike (0-2) count to the state. Second, we will use these win expectancies to compute expectations of future leverage, where the game may evolve in complicated ways, so we expand the space of score differences to -10 to 10.

There is no canonical repository of WE in MLB games incorporating the current ball/strike count into the state, so we must estimate them empirically. As described in the paper, this is a large state space and will lead to poor estimates of WE in some states, even with the large volume of pitch data available to us. We address this data deficiency by simulating 5 million MLB games and then compute WE using those simulated games. The simulation proceeds as follows:

1. As in Tago (2006), we assume MLB games follow a Markov process defined by some game state and transition probabilities to new states.

2. To estimate transition probabilities we need to reduce the game state space from what we have defined above. We define a new limited game state as by the inning part (top or bottom), the score difference between the home and away teams (-10 to 10), the number of outs (0-2), baserunner positions (occupied or not for 1st, 2nd, and 3rd bases), and current ball (0-3) and strike (0-2) count. This removes the inning from the state and assumes games will evolve similarly across innings.

3. Using actual MLB data from our primary estimation sample (2008 – 2018, regular season games), for each possible limited state we compute the probability of transitioning to all potential new limited states.

4. Using these probabilities, we simulate 5 million MLB games from start to finish, as a Markov process in the limited game state. This simulation process evolves through innings. Throughout each simulation, we collect the full game state observed for each pitch and the eventual winner. This information is used to compute the Win Expectancy conditional on a given state as the probability the home team wins in that state across all simulations.

5. Using these WE, we compute our leverage measure for each state as the difference in WE between the case where the state evolves by one strike versus evolving by one ball.

## B.2 Calculation of the expected future leverage metric

The simulation process described above is also the basis for computing expected future leverage. We compute this expectation as follows:

1. For each pitch in the previously simulated games, we assign the value of our leverage metric corresponding to that pitch's game state (inning, inning part, outs, baserunner positions, score difference, ball and strike counts).

2. Again for each pitch, we then compute the sum of this leverage for all future pitches in that half inning of the simulated game. This is the observed future leverage.
3. We compute the average observed future leverage for each possible game state across all simulated games. This is the expected future leverage for that game state.

# C Background on the Game of Baseball

The paper examines the decision-making of officials in the highest level of professional baseball in the United States and Canada. The dynamics we consider are partly driven by the idiosyncratic ruleset of this particular game. For readers unfamiliar with the game, this appendix provides a brief description of rules which are relevant to our analyses followed by a glossary of baseball-specific terms used in the manuscript. Neither is meant to be comprehensive. Major Leage baseball presents a more comprehensive repository of rules and terms at `https://www.mlb.com/glossary`.

## C.1 Summary of Game Structure and Rules

Baseball is a bat-and-ball sport played between two teams of (generally) twenty-five players of which nine are active at any given time in the game. This overview outlines the game's structure, scoring, key player roles, and the critical concept of balls and strikes.

### C.1.1 Game Structure

A baseball game typically consists of nine innings. Innings are divided into two halves – the "top" where the home team is fielding (plays defense) and the visiting team bats (plays offense), followed by the "bottom" where the teams switch position. A team bats until it makes three outs, then the teams switch. If the score is tied after nine innings, extra innings are played until the tie is broken.

### C.1.2 Scoring

A portion of the field, called the infield, consists of four bases in a diamond formation. The batter starts at one base called "home plate". Batting teams attempt to advance players around the bases counter-clockwise. At most one player from the batting team may occupy a particular base at any given time. A player crossing all the bases and returning to home plate scores a run. The fielding team attempts to play batted balls and induce outs of players on the batting team through several methods, including tagging a baserunner with a ball while the runner is not in contact with one of the bases.

### C.1.3 The Batter-Pitcher Interaction

Each play in baseball starts with an interaction between a fielding team player called the "pitcher" and a batting team player called the "batter". The pitcher occupies a position approximately in the middle of the diamond of bases. The pitcher throws the ball toward the batter who is positioned at home plate, trying to prevent the batter from hitting it.

Absent constraints, an obvious strategy for the pitcher to avoid hits would be to throw pitches far away from the batter. To require pitchers to throw pitches with the potential for being hit, baseball defines a "strike zone". This is a polyhedral region which, when viewed from above covers the space directly above home plate. The vertical extent depends on the height and stance of the batter. The bottom of the strike zone begins just below the batter's kneecap and the top is the midpoint of the batter's shoulders and beltline. For any pitch thrown, if the batter elects not to swing, it is ruled a "strike" if the ball passes through any portion of the strike zone and a "ball" otherwise. Batters accruing three strikes are called out via a "strikeout". Batters accruing four balls automatically advance to first base via a "walk", advancing other players already on the bases if another player would be moved onto their base. When all bases are occupied, a walk forces a baserunner back to home plate, scoring a run.

#### C.1.4 Role of the Umpire

Typical MLB games are officiated by four umpires, each having a designated position on the field. One of the four is positioned at home plate and, among other responsibilities, adjudicates all ball-strike determinations. In the event a batter does not swing at a pitch, the home plate umpire observes the pitch, makes the determination of whether it passed through the strike zone, and audibly calls a "strike" if it did or a "ball" otherwise. While many umpire decisions in MLB are reviewable and may be challenged, during the period of our sample, ball-strike decisions were not and the umpire's initial determination is the final ruling. Further, during the period we consider, the home plate umpire did not receive any mechanical assistance or real-time feedback on whether thrown pitches were, in fact, balls or strikes.

### C.2 Glossary of Baseball Terminology

**Ball** A pitch at which the batter does not swing that does not pass through the strike zone. A batter accruing four balls advances to first base on a walk.

**Base** One of four positions arranged in a diamond shape on the field. Baserunners in contact with a base are safe from being called out from being tagged by a batted ball.

**Baserunner** A batter that advances to base either through putting a ball in play or earning a walk. Baserunners attempt to advance through the bases sequentially and score a run if they return to home plate.

**Batter** The player on the batting team who is positioned at home plate and attempts to hit balls thrown by the pitcher using a bat.

**Bottom** The half of an inning occurring second where the home team bats and the away team fields.

**Home Plate** One of the four bases. Batters are positioned at home plate. Baserunners score runs by crossing all of the bases and returning to home plate.

**Inning** A measure of progress of a game. A typical baseball game consists of nine innings where each team alternately takes turns fielding and batting.

**Out** Each half inning continues until the batting team accrues three outs. Batting team players declared out are removed from play. Players may be declared out by striking out or through the play of players on the fielding team.

**Pitch** A ball thrown by the pitcher toward home plate which the batter attempts to hit

**Pitcher** The player on the batting team who is positioned at the pitching mound and throws balls toward home plate where the batter attempts to hit them.

**Run** The unit of scoring in baseball. Runs are scored when a player from the batting team crosses all bases and returns to home plate.

**Strike** A pitch at which the batter does not swing at that passes through the strike zone. A batter accruing three strikes is called out in a strikeout.

**Strike Zone** A polyhedral region covering the space directly above home plate with its bottom defined as the horizontal plane just below the batter's kneecap and the top as a plane at the midpoint of the batter's shoulders and beltline. If a batter elects not to swing at a pitch, it is called a strike if the ball contacts any portion of the strike zone. Otherwise it is called a ball.

**Top** The half of an inning occurring first where the home team fields and the away team bats.

**Umpire** Officials in baseballs games. Typical MLB games use four umpires. One is positioned behind home plate and is solely responsible for adjudicating ball-strike decisions.

# Online Appendix Tables and Figures

Table A.1: Robustness: Omit 9th Inning

| | Primary Spec (1) | Excl. 9th Inning No Lags (2) | Excl. 9th Inning 1 Half Inning Lags (3) | Excl. 9th Inning 2 Half Inning Lags (4) |
|---|---|---|---|---|
| Current Leverage | 37.1577*** | 42.9516*** | 42.9518*** | 42.9501*** |
| | (1.8696) | (2.0743) | (2.0743) | (2.0744) |
| Past Leverage | -1.4971*** | -1.6272*** | -1.6263*** | -1.6251*** |
| | (0.1322) | (0.1470) | (0.1470) | (0.1471) |
| Expected Future Leverage | -3.5497*** | -3.9260*** | -3.9258*** | -3.9206*** |
| | (0.2127) | (0.2418) | (0.2418) | (0.2425) |
| Lag Leverage Half Inning - 1 | | | 0.0427 | 0.0422 |
| | | | (0.1009) | (0.1009) |
| Lag Leverage Half Inning - 2 | | | | -0.0305 |
| | | | | (0.1063) |
| N Pitches | 2,692,669 | 2,428,569 | 2,428,569 | 2,428,569 |
| N Clusters | 26,534 | 26,533 | 26,533 | 26,533 |
| Correct Rate | 0.840 | 0.841 | 0.841 | 0.841 |

Estimates from linear probability model that the umpire makes the correct call for a given pitch. Standard errors clustered at the game level shown in parenthesis. All coefficients and standard errors multiplied by 100 for legibility. Past leverage is the total of current leverage in the current half-inning. Lag leverage is the average of the leverage measure for all ball/strike decisions by the umpire during a previous half-inning. Regressions include game fixed effects. Estimates limited to innings 3-9.

Table A.2: Estimated Bias from Measurement Error in Game Data Leverage

| Minimum Num Games | Variance Signal | Variance Noise | Signal-to-Noise Ratio | Bias Factor b / 03B2b |
|---:|---:|---:|---:|---:|
| 1 | 0.0001764 | 0.0012973 | 0.1196910 | 835.5% |
| 100 | 0.0001533 | 0.0004345 | 0.2607827 | 383.5% |
| 250 | 0.0001390 | 0.0003353 | 0.2930936 | 341.2% |
| 750 | 0.0001169 | 0.0002419 | 0.3258923 | 306.8% |
| 1,000 | 0.0000958 | 0.0002327 | 0.2917219 | 342.8% |
| 2,500 | 0.0000809 | 0.0002148 | 0.2735355 | 365.6% |

Estimates of the magnitude of attenuation bias due to measurement error by using actual game, as opposed to simulated, outcomes to compute leverage. Assumes simulated leverage is the ``true'' measure of leverage. The ``signal'' is this true value ($x$). ``Noise'' is the difference between the GDO leverage measure and the simulated measure for a given pitch ($u$). Under classical measurement error, attenuation bias is proportional to the signal-to-noise ratio $\left(\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2}\right)$ in the probability limit. The Bias Factor is the ratio of the true $b$ (absent attenuation bias) and the estimated $b$ when using the GDO leverage measure assuming classical measurement error. Minimum number of games denotes the minimum number of games on which the GDO leverage measure is based.

Table A.3: Comparison of Estimated Leverage Effects from Different Measures

| Minimum Num Games | GDO Leverage | Simulated Leverage | Bias Factor b / 03B2b |
|---:|---:|---:|---:|
| 1 | -0.528 | 15.216 | -2880.3% |
| 100 | 1.357 | 19.791 | 1457.9% |
| 250 | 3.792 | 23.607 | 622.5% |
| 750 | 9.389 | 47.419 | 505.0% |
| 1,000 | 6.498 | 43.857 | 675.0% |
| 2,500 | 15.347 | 45.336 | 295.4% |

Estimates from linear probability models that the umpire makes the correct call of a given pitch. "Simulated Leverage" estimates computed using our preferred leverage measure from 5 million simulated MLB games. "GDO" (Game Data Only) estimates computed using a leverage measured derived only from actual game data. Standard errors clustered at the game level shown in parenthesis. Attenuation ratio shows the ratio of the estimated coefficients from each model. All coefficients and standard errors multiplied by 100 for legibility. Regressions include only contemporaneous leverage, game, inning, and inning part fixed effects. The first column uses all non-missing observations. Each subsequent column limits to observations where the GDO leverage measure is computed using a minimum of the number of games shown in the column header.

Table A.4: Decomposition of Past Leverage Effects on Umpire Decisions

| | Primary Spec (1) | Calls Only (2) | Call/No Call Decomp (3) |
|---|---|---|---|
| Current Leverage | 37.1575*** | 37.6186*** | 37.8602*** |
| | (1.8696) | (1.8708) | (1.8788) |
| Past Leverage on All Pitches | -1.4971*** | | |
| | (0.1322) | | |
| Past Leverage on Decisions | | -3.2661*** | -2.8802*** |
| | | (0.2773) | (0.3973) |
| Past Leverage on Non-Decisions | | | -0.4308 |
| | | | (0.3150) |
| Expected Future Leverage | -3.5495*** | -3.4466*** | -3.4914*** |
| | (0.2127) | (0.2108) | (0.2133) |
| N Pitches | 2,692,666 | 2,692,666 | 2,692,666 |
| Mean Correct Rate | 0.8403 | 0.8403 | 0.8403 |
| Mean Past Leverage | 0.0634 | 0.1666 | |
| Past Leverage Measure | All Pitches | Decisions | Decisions Non-Decisions |

Estimates from linear probability models that the umpire makes the correct call of a given pitch. Column 1 repeats the primary specification, which computes past leverage as the sum of leverage on all previous pitches in the current half-inning, regardless of whether the umpire made a decision (Past Leverage on All Pitches). Column 2 (Calls only) replaces our preferred past measure with one which accumulates past leverage only for pitches where the umpire was required to make a decision (Past Leverage Calls Only). Column 3 (Calls/No Call Decomp) decomposes our preferred leverage measure into two sums: pitches where the umpire was required to make a decision (Past Leverage all Pitches) and pitches where the umpire did not make a decision (Past Leverage on Non-Decisions).

Table A.5: Effects by Current Score Differential

| | Batting Team Winning (1) | Batting Team Losing (2) | Teams Tied (3) |
|---|---|---|---|
| Current Leverage | 60.4382*** | 30.4999*** | 44.8515*** |
| | (4.8553) | (2.5824) | (3.6563) |
| Past Leverage | -0.9705*** | -2.3813*** | -1.6474*** |
| | (0.2241) | (0.2201) | (0.3089) |
| Expected Future Leverage | -5.7243*** | -3.5698*** | -5.0967*** |
| | (0.7116) | (0.2979) | (0.5983) |

Estimates from linear probability model that the umpire makes the correct call for a given pitch. Standard errors clustered at the game level shown in parenthesis. All coefficients and standard errors multiplied by 100 for legibility. Past leverage is the total of current leverage in the current half-inning. Lag leverage is the average of the leverage measure for all ball/strike decisions by the umpire during a previous half-inning. Regressions include game fixed effects. Estimates limited to innings 3-9.

Table A.6: Effects in Extra Innings

| Effect of Leverage on Correct Calls in Extra Innings | | |
|---|---|---|
| | Primary Spec | Inning 10+ |
| | (1) | (2) |
| Current Leverage | 37.1577*** | 31.9240*** |
| | (1.8696) | (6.0924) |
| Past Leverage | -1.4971*** | -2.1102*** |
| | (0.1322) | (0.5695) |
| Expected Future Leverage | -3.5497*** | -4.2308*** |
| | (0.2127) | (1.0723) |
| N Pitches | 2,692,669 | 60,478 |
| Mean Correct Rate | 0.8403 | 0.8380 |

Estimates from linear probability model that the umpire makes the correct call for a given pitch. Column (1) repeats the estimates from our primary specification, which is estimated using data from innings 2 - 9. Column (2) replaces the estimation sample with ``extra innings'' or innings number 10 or greater. Standard errors clustered at the game level shown in parenthesis. All coefficients and standard errors multiplied by 100 for legibility. Past leverage is the total of current leverage in the current half-inning. Regressions include game fixed effects.

Table A.7: Effects Interacted with Game Attendance

| | Primary Spec (1) | Attendance Interactions (2) |
|---|---|---|
| Current Leverage | 37.1577*** | 31.5676*** |
| | (1.8696) | (5.8383) |
| Past Leverage | -1.4971*** | -0.8899** |
| | (0.1322) | (0.4100) |
| Expected Future Leverage | -3.5497*** | -3.0070*** |
| | (0.2127) | (0.6625) |
| *Interact /w Game Attendance (1000s)* | | |
| Current Leverage | | 0.1812 |
| | | (0.1840) |
| Past Leverage | | -0.0198 |
| | | (0.0129) |
| Expected Future Leverage | | -0.0178 |
| | | (0.0207) |
| N Pitches | 2,692,669 | 2,682,102 |
| Mean Crowd Size (1000s) | | 30.425 |

Estimates from linear probability model that the umpire makes the correct call for a given pitch. Column (1) repeats the estimates from our primary specification. Column (2) interacts leverages measures with the reported game attendance in 1000s. Standard errors clustered at the game level shown in parenthesis. All coefficients and standard errors multiplied by 100 for legibility. Past leverage is the total of current leverage in the current half-inning. Regressions include game fixed effects.