

An inspection game model of the stolen base in baseball: A theory of theft¹

Theodore L. Turocy
School of Economics and CBESS
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
T.Turocy@uea.ac.uk

August 22, 2014

¹The author thanks seminar and conference participants at Texas A&M University, the Stanford Institute for Theoretical Economics, the North American Summer Meeting of the Econometric Society and the World Congress of the Game Theory Society for helpful comments and suggestions on previous drafts of this work. The raw play-by-play data used in this paper are made publicly available on the website of Retrosheet (<http://www.retrosheet.org>). All errors remain the responsibility of the author.

Abstract

This paper models the stolen base play in baseball as a simple inspection game. The model offers equilibrium predictions relating the frequency with which a stolen base play is attempted, and the frequency with which it is successful. Using an extensive play-by-play dataset from 37 Major League Baseball seasons, qualitative and quantitative support is found for the predictions of the model. An exogenous change in the average number of runs scored per game during the period covered by the dataset provides a natural experiment; the equilibrium model predicts the change in the relationship between attempt and success frequencies observed in the data.

Keywords: mixed strategy equilibrium, inspection game, baseball

JEL Classification Numbers: C72, C73

1 Introduction

The analysis of two-player zero-sum games offers perhaps the most compelling prediction for how agents should behave in games, especially among games in which the Nash equilibrium involves randomization. When multiple players randomize in equilibrium, precisely-defined probabilities must be used to ensure players are indifferent. In turn, players whom the equilibrium requires to randomize do not have strong incentives to employ the required probabilities, as they are indifferent among the strategies in the support of their randomization. In the case of two-player zero-sum games, randomize strategies have a more intuitive minimax justification: any deviation from the required probabilities can be exploited systematically by the opponent.

Interactions which are truly zero-sum are perhaps rare. Prominent examples of two-player zero-sum games arise in sports contests between individuals or teams. This paper uses the theory of two-player zero-sum games to develop a simple, inspection game model of the stolen base play in baseball. Baseball players are heterogeneous with respect to their ability to execute this play. The model generates a minimax prediction relating the frequency with which a player attempts the play, and the frequency with which he is successful. The predictions match well with performance derived from a complete play-level dataset from Major League Baseball from 1974 to 2011. Within this time period, there was an exogenous change in the average level of scoring per game. The model makes predictions about how the relationship between attempt and success frequencies should change under the altered conditions, which are consistent with the observed behavior.

The nature and culture of baseball make it a natural laboratory for studying questions of strategy. Baseball proceeds in a series of discrete plays, and the state of the game can be well-summarized by a small number of attributes (the inning, the score, the number of outs, etc.) Baseball also enjoys a culture of detailed record-keeping, which means that transitions among states can be tabulated. Therefore, the ingredients for quantitative study for strategy are present. Lindsey (1977) leveraged these properties to make static comparisons of baseball strategy in a decision-theoretic framework, and Bellman (1977) illustrated how to model the progress of a game of baseball using Markov chains and dynamic programming.

Most previous studies of strategy in baseball, such as that of Lindsey (1977), have taken a decision-theoretic approach, insofar as either offense or defense have a choice to make at any given point in time. This paper incorporates an explicitly game-theoretic element by studying a simultaneous-move interaction. Previous decision-theoretic approaches to the stolen base play predict only a minimum success percentage, and make no prediction about how frequently the play should be attempted.

Evidence generally favorable to the predictions of minimax equilibria have been found in empirical studies in other sports. Walker and Wooders (2001) study service behavior in championship

tennis matches, and cannot reject the hypothesis that servers with the same frequency when serving to the opponent's forehand versus backhand sides. Additionally, they find that professionals seem to do better at choosing their behavior in a serially uncorrelated fashion, compared to behavior reported in laboratory games. Chiappori et al. (2002) consider penalty kicks in professional soccer. As the same striker and goalkeeper will rarely face each other in a penalty kick situation more than once, they study predictions of minimax play in a simple game which are robust to player heterogeneity. Their model generates several qualitative predictions which organize the observed data well.

The richness and structure of sports data also permits testing for evidence that players and teams respond to incentives. In baseball, Bradbury and Drinen (2007) test the hypothesis that pitchers are more willing to risk hitting batters with pitches when they are protected from retaliation due to the use of the designated hitter rule, and Bradbury and Drinen (2008) look for evidence whether pitchers behave differently depending on whether the batter following the current one is stronger or weaker. In association football, Moschini (2010) exploits an exogenous rule change awarding three points for a victory instead of two, and finds that the new rule increases scoring and decreases draws, as intended by the rule designers, consistent with a simple theoretical model.

The paper is organized as follows. Section 2 gives a brief introduction to the stolen base play and develops an idealized model of the play as a two-player, zero-sum inspection game. This model generates three predictions for how stolen base performance should vary across players and across strategic environments. Section 3 takes these predictions to detailed data from 37 seasons of Major League Baseball games, and finds behavior matches the predictions both qualitatively and quantitatively. Section 4 concludes with a discussion of some further open questions in the study of on-field strategy in baseball and other sports.

2 Model

2.1 A brief introduction to the stolen base play

In baseball, two competing teams take turns on offense and defense. While on offense, each team attempts to score runs by advancing team members around a sequence of four bases. The team's turn on offense is terminated when three of its members have been "put out," which can occur by various means. Therefore, outs are a scarce resource for a team. The game is won by the team scoring the most runs after nine innings (i.e., nine turns on offense for each team). Ties are broken by playing further innings, until one team has a lead at the end of an inning.

The players on a baseball team bat in a strict rotation (the "batting order"). As in many bat-and-ball games (such as cricket or rounders), most advancement occurs on batted balls, that is, when

the current batter successfully strikes a pitched ball. However, attempts to advance in baseball are permitted at any time. An attempt to advance a base without the benefit of the ball being batted is called an attempt to “steal” a base. A player successful in stealing a base has advanced one base closer to his ultimate goal of scoring a run; a player who is unsuccessful is put out, costing the team one of its scarce outs.

A stolen base attempt can be thought of as a race. In this race, the offensive player (the runner) runs a distance of about 90 feet. His opponents are members of the defense, the pitcher and the catcher. The pitcher pitches the ball to the catcher, a distance of about 60 feet, and then the catcher relays the ball to the base, a throw of an additional 120 feet. If the runner reaches the base prior to the relay throw from the catcher, the attempt is successful; if not, the attempt fails. While the runner need not wait until the pitcher starts to pitch the ball to begin running, doing so would result in almost certain failure, as the pitcher is permitted instead to throw directly to the base to which the runner is advancing. Therefore, the runner wants to time his departure to closely match the start of the pitcher’s throw to the plate. The choices of the offense and defense are made essentially in ignorance of each other, making a simultaneous-move model the natural choice for describing the interaction.

The interaction between a pitcher and a player known for stealing bases is often described in language suggestive of mixed-strategy equilibrium. Pitchers are encouraged to make the runner’s ability to time his departure more difficult by varying the type, timing, and style of their delivery of pitches; runners, for their part, try to avoid patterns in their behavior, such as always attempting to steal at the first opportunity. Professionals thus perceive some advantage to unpredictability in this setting, which suggests that an equilibrium in mixed strategies should be a feature of an appropriate organizing theory. Finally, the wide availability of data in modern professional sport ensures that common knowledge of the abilities of the relevant players obtains; therefore, it is reasonable to model this interaction as one of perfect information regarding the relevant parameters.

2.2 A formal model

A particular state of a baseball game can be described by a state vector containing, for example, the inning, score, number of outs, and other relevant factors. Suppose that the offense has a runner on first base. At such a point in the progress of the game, the future continuations of the game can be summarized by a vector v , which expresses the probability the team currently on offense will eventually win the game, conditional on the outcome of the interaction to be described next.

The interaction is modeled as a simultaneous-move game. In this game, the offense chooses whether to attempt the stolen base play (strategy S) or not (strategy N). Meanwhile, the defense chooses whether to focus their efforts on trying to put out the batter currently at the plate (strategy

	batter (B)	runner (R)
attempt (S)	$\beta v_S + (1 - \beta)v_F$	$\rho v_S + (1 - \rho)v_F$
no attempt (N)	v_B	v_R

Table 1: A model of the stolen base play as a zero-sum game between the offense and the defense. In the table, the offense is the row chooser, and the defense the column chooser. The cell entries are the payoffs to the offense, measured as the probability the game will eventually be won by the team on offense.

B) or trying to interdict a possible stolen base attempt by focusing on the runner (strategy R). If the stolen base play is attempted, there are two possible outcomes: success, resulting in the runner reaching the next base safely, and failure, resulting in the runner being put out.

The structure of this game is presented in Table 1. It is assumed that both teams seek to maximize the probability of eventually winning the game. Each entry in the table is the probability the team on offense will eventually win the game, conditional on the corresponding strategy profile being chosen. The payoff to the defense is one minus that of the offense.

The vector v of continuation values has four components. The continuation value after a successful attempt is v_S , and v_F is the continuation value after a failed attempt. The continuation values v_B and v_R describe the continuations where the play is not attempted.

The defense's strategy R is their "inspection" strategy. When playing this strategy, only a fraction ρ of stolen base attempts are successful. When B is played, a stolen base attempt is successful with probability $1 \geq \beta > \rho$. The quantity $\beta - \rho$ can be thought of, then, as a measure of the effectiveness of the inspection strategy against the stolen base play.

The analysis assumes the game has a unique equilibrium in which both players randomize. In particular, this requires that $v_R > v_B$. That is, in the contingency in which the offense does not choose to attempt the stolen base, the offense is better off when the defense plays R than when it plays B , and encodes the inspection cost to the defense. This cost derives from the idea that, in order to lower the chance of a stolen base play's success, the defense must modify their pitching approach to the current batter. If there were no runner on base, the pitcher and batter can be thought of as engaging in their own zero-sum game, not modeled explicitly here, with the pitcher choosing the type and location of pitch to throw, and the batter forming expectations about the pitch. Now, with the baserunner on first base, let the strategy B in the game corresponds to following the optimal strategy against the batter as if there were no runner on, and R corresponds to following some modified pitching strategy to defend against the runner. In the latter case, it must be that the defense is no longer using their minimax strategy against the batter, and so the batter will perform better. Smith (1980), for example, highlights the role of this effect, writing

Any consideration of base running must include indirect and often subtle effects....

The greatest of these indirect effects is of course the intangible of upsetting the pitcher by diverting his attention from the batter.

When the equilibrium is in mixed strategies, the equations for the equilibrium probabilities of attempting the stolen base play, p_S^* , and for focusing attention on the batter, p_B^* , are

$$p_S^*(\rho; \beta) = \frac{v_R - v_B}{(\beta - \rho)(v_S - v_F) + (v_R - v_B)} \quad (1)$$

$$p_B^*(\rho; \beta) = \frac{v_R - [\rho v_S + (1 - \rho)v_F]}{(\beta - \rho)(v_S - v_F) + (v_R - v_B)} \quad (2)$$

In taking the model to the data, the focus will be on players for whom attempting the stolen base play is a salient activity. For this subset of players, received wisdom would suggest that β will be close to unity; they are players who tend to be quick afoot, and would be able to be successful almost every time if countermeasures were not deployed.¹ The expectation is that heterogeneity across players arises primarily in the parameter ρ . For example, former Oakland Athletics coach Ron Washington has said,

“A base stealer is a guy who when everyone in the ... yard know he gonna get the bag, he gets the bag.”²

A more precise (if admittedly less colorful) way of phrasing this is that a player who is effective at the stolen base play will be successful a high fraction of the time, even when the play is anticipated. In the model’s formulation, this corresponds to a relatively large value of ρ .

2.3 Testable implications

The parameters ρ and β are not directly observable. This subsection derives three testable implications outlining qualitative and quantitative predictions about how stolen base play behavior should change across players and across strategic environments.

Write the percentage of stolen base attempts which are successful as

$$\pi^*(\rho) \equiv \beta p_B^*(\rho) + \rho [1 - p_B^*(\rho)] . \quad (3)$$

¹In 1974-1975, the Oakland Athletics employed Herb Washington, a world-class sprinter with no advanced baseball playing experience, as a pinch-runner, in part to exploit the possibility of the stolen base play at crucial situations in a game.

²Quoted in *Moneyball: The Art of Winning an Unfair Game*, by Michael Lewis, W. W. Norton & Company, 2003, page 265.

Prediction 1. *The relationship between the frequency π with which a runner attempts a stolen base, and the frequency p_S with which his attempts are successful, satisfy the affine relationship*

$$\pi = \frac{v_B - v_F}{v_S - v_F} + \left[\beta - \frac{v_B - v_F}{v_S - v_F} \right] p_S. \quad (4)$$

Justification. Begin by substituting (2) into (3),

$$\pi^*(\rho) = \frac{\beta \{v_R - [\rho v_S + (1 - \rho)v_F]\} + \rho \{\beta v_S + (1 - \beta)v_F - v_B\}}{(\beta - \rho)(v_S - v_F) + (v_R - v_B)} \quad (5)$$

$$= \frac{\beta(v_R - v_B) + (\beta - \rho)(v_B - v_F)}{(\beta - \rho)(v_S - v_F) + (v_R - v_B)}. \quad (6)$$

$$= \beta p_S^*(\rho) + \frac{(\beta - \rho)(v_B - v_F)}{(\beta - \rho)(v_S - v_F) + (v_R - v_B)}. \quad (7)$$

Equation (1) can be rearranged to

$$\beta - \rho = \frac{1 - p_S^*(\rho)}{p_S^*(\rho)} \times \frac{v_R - v_B}{v_S - v_F}. \quad (8)$$

Substituting this into (7),

$$\pi^*(\rho) = \beta p_S^*(\rho) + \frac{1 - p_S^*(\rho)}{p_S^*(\rho)} \times \frac{v_R - v_B}{v_S - v_F} \times \frac{v_B - v_F}{(\beta - \rho)(v_S - v_F) + (v_R - v_B)}. \quad (9)$$

$$= \beta p_S^*(\rho) + (1 - p_S^*(\rho)) \times \frac{v_R - v_B}{v_S - v_F} \times \frac{v_B - v_F}{v_R - v_B} \quad (10)$$

$$= \beta p_S^*(\rho) + (1 - p_S^*(\rho)) \times \frac{v_B - v_F}{v_S - v_F} \quad (11)$$

$$= \frac{v_B - v_F}{v_S - v_F} + \left[q - \frac{v_B - v_F}{v_S - v_F} \right] p_S^*(\rho). \quad (12)$$

□

The values of the slope and intercept in (12) depend on the current situation in a game, as well as the broader environment. In the history of Major League Baseball, the average number of runs scored per game has varied significantly. Intuitively, when runs are easy to come by, the value of a single run towards increasing a team's chances of victory becomes less. This in turn increases the value of the ratio $\frac{v_B - v_F}{v_S - v_F}$, and gives a second testable prediction of the model.

Prediction 2. *In seasons in which runs scored per game are higher, the relationship (12) will have a larger intercept and flatter slope than in seasons in which runs per game are lower.*

Justification. Baseball play-by-play data is sufficiently rich that the probability a team eventually wins a game, starting from any given game situation, can be estimated using various approaches.

Consider a game that is in the first inning, with no outs and a runner on first base only. Suppose the expected number of runs scored in the game by each team is 4.0. Using a simulation approach,³ v_S is estimated to be 0.5621 and v_F to be 0.4786. Because the defensive strategy is not easily observable, the expected win probability from the game state itself is $p_B v_B + (1 - p_B) v_R$, which is estimated at 0.5364; this is therefore an upper bound on v_B . Given these estimates, it follows that

$$\frac{v_B - v_F}{v_S - v_F} < \frac{0.5364 - 0.4786}{0.5621 - 0.4786} = 0.692. \quad (13)$$

Now instead consider the same situation, except in a setting where the expected number of runs per game is 5.0 per team. Then, the corresponding simulation estimates are $v_S = 0.5567$, $v_F = 0.4771$, and $v_B < 0.5346$, and so

$$\frac{v_B - v_F}{v_S - v_F} < \frac{0.5346 - 0.4771}{0.5567 - 0.4771} = 0.722. \quad (14)$$

Also, observe that rather small changes in v_B , holding the other values constant, moves the value of the ratio substantially. So even if the inspection cost $v_B - v_R$ is small, the resulting values of the ratios will be several percentage points below these upper bounds.

Other methodologies yield comparable estimates for this ratio. In the first inning of a game, maximizing the expected number of runs scored in the inning is approximate the same as maximizing the probability of winning, insofar as endgame effects are negligible. Thorn and Palmer (1984) present a table of the expected number of future runs in an inning as a function of the current number of outs and location of runners, using a different simulation model based on empirical data from the seasons between 1961 and 1977. Their table gives $v_S = 1.068$, $v_F = 0.249$, and $v_B < 0.783$, and therefore

$$\frac{v_B - v_F}{v_S - v_F} < \frac{0.783 - 0.249}{1.068 - 0.249} = 0.652. \quad (15)$$

The period from which Thorn and Palmer sample spans the lowest-scoring period in Major League Baseball history,⁴ which is consistent with the intuition that the ratio's value will be lower in environments where expected scoring is lower. $\square$

The model incorporates the inspection cost to the defense in the assumption that $v_R > v_B$, justified by the assertion that actions which lower the chance of a successful stolen base attempt result in an improved performance by the batter. While the extant data do not provide reliable

³The values cited here are provided by The Hardball Times website's Win Probability Inquirer, at http://www.hardballtimes.com/thtstats/other/wpa_inquirer.php

⁴Run scoring reached its nadir in 1968, with only 3.42 runs scored per game. Most seasons between 1963 and 1972 saw fewer than 4 runs per game scored. These seasons predate the sample used in the analysis.

indicators of the conduct of the defense, nevertheless there is a qualitative prediction about how the performance of batters should change.

Prediction 3. *Batters should in general benefit from having a runner on first base. Further, there should be a positive relationship between the size of this batter performance benefit, and the frequency with which a runner on first attempts to steal bases.*

Justification. In the long run, a batter's observed performance will be the weighted average of his performance when strategy B is played and his performance when strategy R is played. Both $p_S^*(\rho)$ and $p_B^*(rho)$ are increasing in ρ . Therefore, the relative frequency of the outcome (N, R) versus that of (N, B) increases as p_S increases, and so batter performance, conditional on no stolen base attempt, should correlate positively with the frequency with which stolen bases are attempted. $\square$

3 Empirical evidence

3.1 Predictions 1 and 2: Relating attempt and success frequencies

Based on (12), consider the linear regression model

$$\pi = \alpha_0 + \alpha_1 \times p_S + \varepsilon. \quad (16)$$

Let OPP_i denote the number of opportunities player i has to attempt the stolen base play, ATT_i denote the number of stolen base plays attempted by that player, and SUC_i the number of attempts which are successful. Then, the regression model can be expressed as

$$\frac{SUC_i}{ATT_i} = \alpha_0 + \alpha_1 \times \frac{ATT_i}{OPS_i} + \varepsilon \quad (17)$$

$$SUC_i = \alpha_0(ATT_i \times OPP_i) + \alpha_1 \times ATT_i^2 + (ATT_i \times OPP_i) \times \varepsilon. \quad (18)$$

Equation (18) is estimated using weighted least squares, to account for the fact that players differ both in the number of opportunities in the dataset, as well as the frequency with which attempts are made.

Complete play-by-play data is available for all Major League Baseball games from 1974 through 2011, inclusive. “Play-by-play” means that the dataset identifies all events that change the number of outs, or the configuration of baserunners. Specifically, the outcomes of the times at bat of all batters are recorded, as well as the timing and outcome of stolen base plays.

The predicted values for α_0 and α_1 depend on the vector of continuation values v . These values in turn depend on the current situation in the game, so simple aggregation of data across all situations would not be valid, as the strategic incentives in different game situations will differ.

The sample is therefore taken from one game situation, when the game is in the first inning, the score is tied, there are no outs in the inning, and there is a runner on first base only. This selection has two appealing properties:

1. It is a common situation. All games begin tied, and the first batter in an inning reaches first base approximately 30% of the time.
2. The same players are frequently involved. The order in which players bat in a baseball team's batting order tends to be stable; most players have a customary slot in the lineup in which they appear in most of the games they play. Furthermore, many players who attempt stolen bases most frequently tend to bat in the first slot in the batting order, implying that they will be on first base in this situation many times over the course of a season.

An observation then consists of the total number of opportunities, attempts, and successes for a player within a given season. Attention is restricted to full-time regular players, defined as those who have $OPP_i \geq 20$.

The predicted values of α_0 and α_1 also depend on the run-scoring environment. Here the play-by-play data span a step-level event which offers a convenient natural experiment. Historically, the number of runs scored per game in Major League Baseball has fluctuated due to variations in playing conditions, equipment, playing styles, and other factors. In the early 1990s, a step-level change in the number of runs scored per game occurred. In 1992, 4.12 runs were scored per game. In 1993, this increased to 4.60; on through 2009, runs per game were at least 4.59, peaking at 5.14 in 2000. Several reasons for this regime change have been proposed, including the addition of four new teams (two in 1993 and two in 1998), the opening of stadia more favorable to offense (including a stadium in Denver, the first stadium at altitude), and an alleged increase in the use of performance-enhancing drugs, including steroids.⁵ Significantly, however, all of these explanations do not bear specifically on the strategic interaction of the stolen base play, and therefore the change in run scoring provides an exogenously-created natural experiment. Equation (12) makes a qualitative prediction for how α_0 and α_1 should change under this shift. In a higher-scoring environment, individual runs become less valuable, increasing the required breakeven success probability. This implies that α_0 should increase, and α_1 should decrease.

For the period 1974-1992, there are 289 player-seasons meeting the criteria; for 1993-2011, there are 336 player-seasons. A scatterplot of the data is presented in Figure 1, with graph elements in gray corresponding to 1974-1992 and those in black to 1993-2011. The figure also includes a weighted local polynomial regression fit for each subsample. Qualitatively, the local polynomial fits match the predictions. Attempt and success frequencies have a positive relationship. The

⁵De Vany (2011) offers an analysis of the distribution of home runs over time in baseball, motivated by the steroid question.

success frequency of players who attempt rarely is in the expected range, and is higher in 1993-2011. The relationship between attempt and success frequency is flatter in 1993-2011 than in 1974-1992.

More formally, the coefficient estimates on model (18) from weighted least squares linear regression are, for 1974-1992,⁶

$$SUC_i = -6.006 + 0.609(ATT_i \times OPP_i) + 0.324(ATT_i)^2 \quad (19)$$

(8.447) (0.037) (0.051)

and for 1993-2011,

$$SUC_i = -2.378 + 0.648(ATT_i \times OPP_i) + 0.237(ATT_i)^2 \quad (20)$$

(6.330) (0.045) (0.087)

The parameter estimates are in line with the predictions of the model:

- Equation (18) requires the intercept term to be zero; the estimated intercepts are not significantly different from zero.
- The estimated magnitudes of α_0 are slightly less than the breakeven percentages computed in the previous section.
- The model predicts $\alpha_0 + \alpha_1 = \beta < 1$. In both equations, $\alpha_0 + \alpha_1 \approx 0.9$, which would correspond to a stolen base play being successful 90% of the time even when strategy B is employed by the defense. No individual player in history has had a success rate of over 90% with more than 50 attempts, so this is a plausible range for β .
- A Chow test rejects the null hypothesis that the parameter estimates are the same between the two time periods (F -statistic 12.3, p -value $\approx 10^{-7}$). Therefore, the comparative statics predictions of the model in response to the exogenous change in the strategic environment are found in the data.

3.2 Prediction 3: Effects on batter performance

The model predicts that an increase in the baserunner's stealing skills results in increased attention paid to the runner by the defense. Therefore, a batter should enjoy better performance, the greater the threat of the stolen base play created by the runner on first base.

A test of Prediction 3 requires an operationalization of how to measure the performance of a batter. Research in both the academic and applied baseball analytics literatures has demonstrated

⁶In all regression equations, standard errors are in parentheses.

that the performance of a baseball team in producing runs can be predicted with substantial accuracy.⁷ One measure is weighted on base-average (wOBA), proposed by Tango et al. (2007), which constructs a performance measure based on outcomes of a batter's time at bat, which correlate strongly at the team level with run-scoring:⁸

$$wOBA = \frac{0.72 \cdot BB + 0.75 \cdot HP + 0.90 \cdot 1B + 0.92 \cdot RE + 1.24 \cdot 2B + 1.56 \cdot 3B + 1.95 \cdot HR}{PA} \quad (21)$$

Let i denote a batter, and j denote a runner, within a given season. For each (i, j) pair, construct the wOBA of batter i in situations in which runner j was the runner on first base, and the other bases were empty; call this $wOBA_{ij}$. This is compared to batter i 's baseline performance, which is measured as his performance in situations in where there are no runners on base, as in this situations there are no inspection game considerations; call this $wOBA_{i\emptyset}$. The performance increase (or decrease) enjoyed by batter i when runner j is present on base is then $\Delta wOBA_{ij} \equiv wOBA_{ij} - wOBA_{i\emptyset}$.

The model predicts that wOBA should be higher when a player with a higher frequency of stolen base attempts is the runner at first base. For each (i, j) pair, let p_{ij} be the probability of a stolen base attempt occurring in the situation in which i is the batter and j is the runner on first base, and other bases are empty.

Again to focus on full-time regular players, the analysis considers (i, j) pairs such that there are at least 30 observations of the pair in the required game situation; there are 5340 such pairs in the sample. Linear regression of $\Delta wOBA$ on p , weighted by the number of plate appearances involving the batter in the sample, gives the regression equation

$$\Delta wOBA_{ij} = \underset{(0.0016)}{0.012} + \underset{(0.011)}{0.035} p_{ij}. \quad (22)$$

Both coefficients are significantly different from zero at the 99% significance level. This is consistent with the two statements in Prediction 3. Having a runner on first base improves batting performance. Performance is enhanced further, the greater the threat of a stolen base by the runner on first. Even though the play-by-play data lacks the granularity for the strategic conduct of the defense to be observed directly, this result provides indirect evidence that the defense, in fact, does change their approach to pitching to the batter in response to the magnitude of the stolen base threat on first base, as game-theoretic logic predicts.

Techniques from Tango et al. (2007) permit a rough translation of the value of this effect in

⁷See e.g. Turocy (2005) for a discussion.

⁸In this formula, BB = bases on balls, HP = times hit by pitch, 1B = singles, 2B = doubles, 3B = triples, HR = home runs, RE = times reached on error, and PA = total number of appearances at bat.

terms of wins and losses for a team. A player with an attempt frequency of 0.5, among the elites in the dataset, would generate a $\Delta wOBA$ of 0.0295 in this situation. A full-time player would be the baserunner on first in such a situation roughly 100 times per season. Tango et al. (2007) converts $wOBA$ to changes in runs scored by the relationship $\Delta RUNS = \frac{\Delta wOBA}{1.15} \times PA = \frac{0.0295}{1.15} \times 100 = 2.57$. To give this figure context, a team has about 5000 plate appearances per season, and the highest-scoring team in a league will typically outscore an average team by about 100 runs. If this effect were present for all 5000 plate appearances, a team would score about 128 more runs in a season. That is, by this rough calculation, a player who is a significant threat to attempt the stolen base play improves the performance of the batter by roughly the difference between the best offensive team and an average offensive team in a league. So the potential size of the effect is significant in rate terms, but because it affects a relatively small percentage of plate appearances in a team's season, the total size of the effect is small relative to a player's other contributions to a team's offense.

4 Conclusion

Minimax equilibrium in a stylized inspection game organizes the observed strategic conduct in stolen base attempts in Major League Baseball. The relationship between the frequency a player attempts the play, and the frequency with which he is successful, is in line with theoretical predictions. When there is an exogenous environmental shift due to increased scoring of runs, this relationship changes consistent with the theory's predictions. Finally, evidence that the defense's strategy changes qualitatively in agreement with the theory is found by examining the increase in batter performance when a player who is a stolen base threat is on base.

These results continue a line of literature including Walker and Wooders (2001) and Chiappori et al. (2002) exhibiting sports interactions where minimax equilibrium organizes observed behavior well. In one sense, this seems almost inevitable. These interactions involve highly-motivated professionals who have access to a wealth of information about the interaction. Even in such settings, however, systematic deviations from optimal behavior can be found. Romer (2006) found that coaches in professional American football are systematically too conservative on attempting to convert fourth-down plays, even though they enjoy access in principle to decades' worth of data on the risks and rewards of such conversion attempts. Therefore, the Romer result cannot be attributed to a lack of relevant data. More probably, it is the artefact of coaches pursuing a more complex objective function, one in which regret from failure on such a play is particularly salient. Such regret could indeed be quite rational, if failure from making the "unorthodox" choice results in second-guessing by fans (or the loss of the coach's job from dissatisfied team owners).

The history of baseball may offer an opportunity to test whether common knowledge of the

parameters of the game is indeed important for underpinning the behavior predicted in equilibrium. Historical play-level data is slowly becoming available from the early days of baseball history, dating back into the nineteenth century. In addition to predating computer analysis, television, and radio, the institutional structure of baseball was less rigidly defined than today. In modern professional sports, including baseball, no player appears in top-class competition without a track record of performance dating back several years. This is increasingly less true in baseball as one moves back into time; even into the 1930s and 1940s, a Major League Baseball team might be confronted with opposing players about whom they knew next to nothing. At present this historical dataset remains too incomplete to carry out an analysis along the lines of this paper. One might expect that the tidy relationships observed in the data on stolen base plays in the modern game disappear when looking far enough back in time to an environment where imperfect information would be the norm.

References

- R. Bellman. Dynamic programming and Markovian decision processes, with application to baseball. In *Optimal Strategies in Sports*, pages 77–85. Elsevier-North Holland, New York, 1977.
- J. C. Bradbury and D. J. Drinen. Crime and punishment in Major League Baseball: The case of the designated hitter and hit batters. *Economic Inquiry*, 45:131–144, 2007.
- J. C. Bradbury and D. J. Drinen. Pigou at the plate: Externalities in Major League Baseball. *Journal of Sports Economics*, 9:211–224, 2008.
- P.-A. Chiappori, S. Levitt, and T. Groseclose. Testing mixed strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*, 92:1138–1151, 2002.
- A. De Vany. Steroids and home runs. *Economic Inquiry*, 49:489–511, 2011.
- G. P. Lindsey. A scientific approach to strategy in baseball. In *Optimal Strategies in Sports*, pages 1–30. Elsevier-North Holland, New York, 1977.
- G. Moschini. Incentives and outcomes in a strategic setting: The 3-points-for-a-win system in soccer. *Economic Inquiry*, 48:65–79, 2010.
- D. Romer. Do firms maximize? Evidence from professional football. *Journal of Political Economy*, 114:340–365, 2006.
- D. W. Smith. Maury Wills and the value of a stolen base. *Baseball Research Journal*, 1980.
- T. Tango, M. Lichtman, and A. Dolphin. *The Book: Playing the Percentages in Baseball*. Potomac Books, Washington, 2007.

- J. Thorn and P. Palmer. *The Hidden Game of Baseball*. Doubleday & Company, Inc., Garden City NY, 1984.
- T. L. Turocy. Offensive performance, omitted variables, and the value of speed in baseball. *Economics Letters*, 89:283–286, 2005.
- M. Walker and J. Wooders. Minimax play at Wimbledon. *American Economic Review*, 91:1521–1538, 2001.

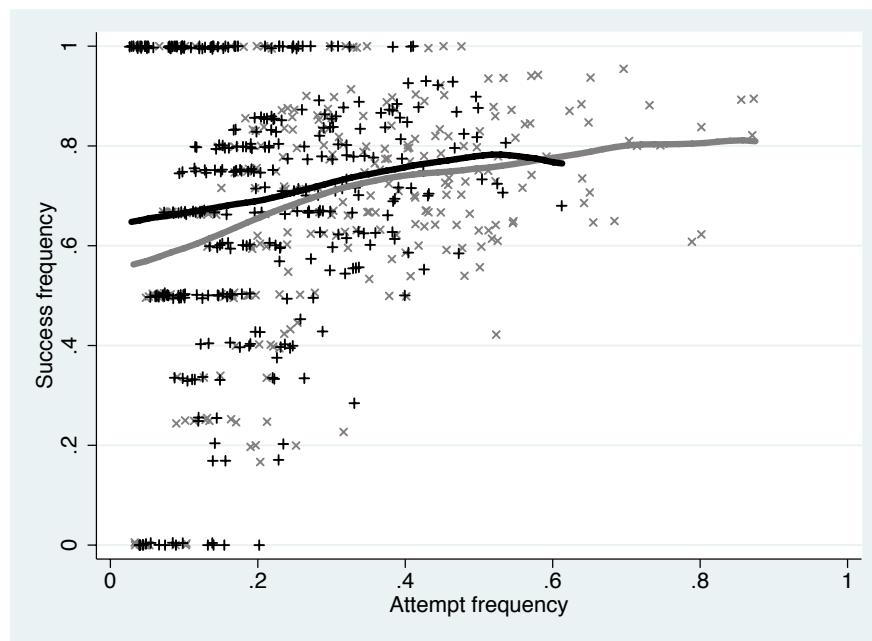


Figure 1: Scatterplot of frequency of attempt of stolen base play versus percentage of attempts successful. Each point represents one runner in one season, with a minimum requirement of having 20 opportunities (as defined by the game situation in the text). Points in grey represent observations in 1974-1992; points in black represent observations in 1993-2011. The solid lines are local polynomial fits to the data (STATA `lpoly`, Epanechnikov kernel, bandwidth 0.08).