Offensive Performance, Omitted Variables, and the Value of Speed in Baseball

by Theodore L. Turocy

Department of Economics
Texas A&M University
College Station TX 77843
Phone: 979.862.8082
Fax: 979.847.8757

Email: turocy@econmail.tamu.edu

December 19, 2003

Abstract

This note considers the problem of estimating the marginal products of offensive events towards a baseball team's objective of scoring runs. Regression techniques on official statistics give a positive marginal product for a stolen base attempt, which is inconsistent with the theory of mixed strategy Nash equilibrium. Augmenting the specification of the production function to include other productive qualities of foot-speed restores estimates consistent with equilibrium theory.

Keywords: Omitted variable bias, mixed strategies, equilibrium test, baseball.
JEL Classification Numbers: C72, C50, L83

1 Introduction

Economists believe that certain basic principles organize a wide range of human activity. This belief is best put to the test in environments where the assumptions underlying theory are most likely to be satisfied. In this spirit, Walker and Wooders [11] investigated the predictions of mixed-strategy Nash equilibrium in the context of play during Grand Slam tennis finals. The assumptions underlying the theory, such as common knowledge of the game, plausibly hold in this environment, and, moreover, the extensive structure and payoffs of the game are well-known to players and modelers alike.

This observation generalizes across professional sports.[1] Sports are games with explicit, well-defined rules, and, at the highest levels, are played with mutual understanding of the strengths and weaknesses of the players involved. Not all sports provide an opportunity

This document has been produced using GNU T_EX_{MACS} (see http://www.texmacs.org).

1. The suggestion of using sports as an analytical testbed dates at least to Mottley [8].

for as clean a dataset as tennis afforded Walker and Wooders; for example, CHIAPPORI ET AL [4], investigating penalty kick behavior in professional soccer, are limited to testing predictions of the theory that survive the presence of player heterogeneity. Despite such limitations, the structure of sports games still presents the economist with opportunities to study highly motivated agents in environments where rules and incentives are well-defined and observable.

This paper pursues this line of investigation by considering the strategic choice of attempting a stolen base in a baseball game. The objective of a baseball team is to win games by scoring more runs than they allow their opponents to score. Generally, a baseball team scores runs by advancing around the bases on batted balls. However, players on base are permitted to attempt to advance at any time the ball is in play; when they attempt to do so when the batter has not hit the ball, they are said to attempt to steal a base.

Undertaking this play is strategic, in the sense that players are not compelled by the rules to make such attempts. Empirically, almost all players are observed to attempt to steal bases with positive probability. Players (and teams) who attempt stolen bases frequently are almost always players (and teams) who carry a reputation for being fast; in fact, it seems clear that footspeed is an important ingredient in being able to reach the next base before the defense is able to put the player out.

The discrete structure of the game permits the progress a team makes towards the goal of scoring runs to be quantified, for example, by counting the number of safe hits, home runs, walks, and so forth a player or team accumulates. Although the number of runs a team actually scores depends on the ordering of these events, a remarkable regularity is that a large percentage of the variation in the output of runs across teams can be explained using simple linear models based on statistics which have been kept as part of baseball's historical record for more than a century.

Indeed, there is a literature, surveyed in Section 2, on estimating parameters for these models. This conversation has occurred primarily outside the economics literature. The economist's contribution to this pursuit is to intepret the results. Supposing a simultaneous-move game model of the offense-defense interaction, the observation that almost all baserunners attempt with positive probability implies that the net value of an attempt, relative to not attempting, should be zero. However, different methodologies for estimating parameters yield results which have different economic implications. In particular, least-squares regression using team-season data yield estimates which imply a positive marginal value for an additional attempted steal. This would be inconsistent with equilibrium theory in this setting.

As noted, players who attempt many stolen bases are generally considered to be the fastest players. Footspeed contributes in other ways to a baseball team's offense. A fast runner may advance a base safely when a slower one would be put out. Also, a fast runner is more likely to attempt to advance further on batted balls than a slower one, pressuring the defense into making more misplays. These are statistics which do not appear in the historical record, and have therefore been omitted by necessity from the cited literature.

In Section 3, the regression specification is augmented with an additional set of explanatory variables which are plausibly correlated with a team's overall speed. These variables, not included in the historical record of baseball, are gleaned from an extensive dataset of play-by-play data from all Major League Baseball games between 1974 and 1992. With these additional controls for team speed added, parameter estimates are recovered which are consistent with an equilibrium theory of stolen base behavior. Section 4 concludes the paper by summarizing some implications for understanding the role of elite basestealers in the context of a team's offensive production.

2 Estimating the production function

The task at hand is to correctly estimate the coefficients of a linear approximation to the production function for producing runs. While the true production function mapping events in a baseball game to runs is not linear, the curvature is gentle, and negligible within the subset of the sample space in which teams' aggregate statistics are observed to lie.[2] The coefficients of the function, then, are estimates of the marginal contribution of an additional offensive event to the production of runs. Table 1 presents estimates of the coefficients of the production function for five batting events, singles (1B), doubles (2B), triples (3B), home runs (HR), and walks (BB), as well as a successful stolen base (SB) and a time caught stealing (CS).

Source	Data	1B	2B	3B	HR	BB	SB	CS
Bennett-Flueck (regression)	1969-77 $N = 218$	0.50 (.026)	0.73 (.073)	1.27 (.164)	1.45 (.055)	0.35 (.022)	0.13 (.041)	N/A
Berry (regression)	1990-98 $N = 248$	0.49 (.029)	0.72 (.072)	1.14 (.190)	1.51 (.053)	0.34 (.024)	0.26 (.054)	-0.14 (.140)
Albert-Bennett (regression)	1954-99 $N = 1078$	0.52 (.012)	0.66 (.023)	1.17 (.067)	1.49 (.024)	0.35 (.010)	0.19 (.023)	-0.11 (.061)
Lindsey (play-by-play)	1959-60	0.46	0.82	1.07	1.42	0.33	N/A	N/A
Palmer (simulation)	1961-77	0.45	0.77	1.00	1.42	0.33	0.19	-0.32

Table 1. Average run values for different types of events, as reported by different authors. For the regression approaches, the numbers in parentheses are standard errors for the coefficients, and N is the number of team-seasons in the regression.

These coefficients were determined using various methodologies and data sets. The first three rows in Table 1 were produced by computing OLS estimates of runs as a linear function of offensive events, reported in Bennett and Flueck [2], Berry [3], and Albert and Bennett [1], respectively. The three specifications differ in what additional regressors are present in the model; nevertheless, the estimated marginal contributions are robust across the different datasets and specifications.

2. Discussion of the shape of the "true" production function can be found in Albert and Bennett [1].

The fourth and fifth rows represent a micro-level approach to the estimation. LINDSEY [7] analyzed play-by-play data obtained by observing a sample of Major League Baseball games in 1959 and 1960. From this data, he constructed a table of the expected number of runs scored in the remainder of a team's inning at bat as a function of the current number of outs and the bases occupied by runners.[3] The value of an offensive event such as a single was then calculated to be the expected change in this number of subsequent runs scored, taking into account the distribution of baserunners' advancement on the play.

THORN AND PALMER [9] report the results of a similar calculation. Instead of using actual play-by-play data, Palmer created a Markov chain simulation model of each team's offense using its aggregate statistics. In the model, a team would hit, for example, a single with the same probability it hit singles in the actual season. Lacking play-by-play data, features of the model such as the number of bases baserunners would advance on a single were simulated using overall (i.e., not team-specific) frequencies. Based on the simulated play-by-play, Palmer produced his coefficient estimates using an a technique analogous to Lindsey's.

Four of the five authors cited also report estimates for the marginal value of a successful stolen base and of a time caught stealing.[4] Unlike the coefficients on batting events, the estimates produced by regression on observed team-level statistics have substantially different implications from the estimates produced by Palmer's simulation approach. This difference can be seen by pursuing a micro-level analysis based on the methods of Lindsey and Palmer. Table 2 provides the expected future runs scored in an inning, as a function of the number of outs and the currently occupied bases, generated by Palmer's simulation.

	0 out	1 out	2 out
Empty	0.454	0.249	0.095
1st	0.783	0.478	0.209
2nd	1.068	0.699	0.348
3rd	1.277	0.897	0.382
1st and 2nd	1.380	0.888	0.457
1st and 3rd	1.639	1.088	0.494
2nd and 3rd	1.946	1.371	0.661
Loaded	2.254	1.546	0.798

Table 2. Expected future runs scored in an inning, as a function of the number of outs and the currently occupied bases. This table is reproduced from THORN AND PALMER [9] based on the results of Palmer's simulation on the 1961-77 dataset.

3. Lindsey's table resembles Table 2, augmented by additional information about the distribution of runs scored.

4. The regression reported in Bennett and Flueck's is their Expected Run Production model, which was selected based on an adjusted R^2 criterion; this regression did not include times caught stealing due to its high correlation with successful stolen bases.

If one were to assume for the moment that these data were generated by the interaction of offenses and defenses with the goal of maximizing (or minimizing, for the defense) expected runs, and that this interaction is in an equilibrium state, then the probability a stolen base is successful must be such that the net benefit of a stolen base attempt is zero, relative to not attempting. For example, with a runner on first and no outs, the expected number of future runs in the inning is 0.783. Assume for simplicity that the two outcomes of attempting a steal of second are success, resulting in a runner on second with no outs, and failure, resulting in no runners on an one out. Letting π be the probability a stolen base attempt is successful, equilibrium would imply that

$$1.068\pi + 0.249(1 - \pi) = 0.783.$$

This equation is solved by $\pi = .652$. Parallel calculations give an equilibrium $\pi = .634$ for the case with a runner on first and one out, and $\pi = .601$ for a runner on first and two out.

The success percentage which results in a net gain of zero from the marginal stolen base attempt using Palmer's estimates is $\frac{.32}{.51} = .627$, which is comparable to the values of π determined above. On the other hand, the Albert and Bennett coefficient estimates imply a net gain of zero occurring with a success percentage of $\frac{.11}{.30} = .367$, and Berry's give an implied success percentage of $\frac{.14}{.40} = .35$. From 1954 through 1999, the range of Albert and Bennett's data, only three clubs posted a stolen base success percentage lower than .367: the Washington Senators in 1957 (0.255) and again in 1958 (0.349), and the 1978 Toronto Blue Jays (0.350). The regression estimates would imply that teams are stealing far too infrequently, as they would still be enjoying substantial gains from stolen base attempts at the margin. Palmer's estimates are consistent with the predictions of theory.

3 Controlling for speed

The unsatisfactory economic implications of the SB and CS coefficients from regression analysis arise because the specification of the production function omits inputs related to speed. Prolific basestealers are almost always fast, and it is reasonable to expect that speed is helpful in avoiding outs and producing runs. The productive effects of speed are captured in the regression analyses in the coefficients on the speed-related statistics 3B, SB, and CS. On the other hand, the design of Palmer's simulation represents essentially a perfect control for all other effects of speed, since baserunner advancement in the simulation was implemented using overall, not team-specific, frequencies.

This suggests that the inclusion of other productive (non-strategic) factors related to speed into the team-level regressions should result in coefficient estimates for SB and CS which are comparable in economic implication to those of Palmer. To improve the specification of the production function, this analysis turns to the database of play-by-play data collected by Retrosheet[5] to tabulate frequencies of other offensive activities which are plausibly correlated with speed.

The official records of baseball lump all stolen bases and times caught stealing together, regardless of whether it is an attempted steal of second, third, or home. Table 2 implies that the equilibrium success percentages for steals of third are substantially different than for steals of second. Under the equilibrium assumption, with a runner on second only and no out, the success percentage for a steal of third base is .797. With one out, the percentage is .753; with two out, the percentage sharply increases to .911. These imply that, while it may be optimal for typical baserunners to attempt the occasional steal of second base, only elite baserunners will attempt steals of third with significant frequency.[6] Using the Retrosheet data, basestealing statistics are disaggregated by the base being attempted. SB2 and CS2 denote successful and unsuccessful steals of second base, and SB3 and CS3 successful and unsuccessful steals of third base.

Also from the play-by-play data, the number of times a baserunner advanced on an error by the defense is tabulated. Faster baserunners will attempt to advance further on plays, forcing fielders to make throws to attempt to put them out. Under pressure from a baserunner's superior speed, errant throws may be more likely.

In addition, the specification includes times grounded into double play (GDP). GDP has a fairly strong negative correlation with stolen bases and caught stealing for two reasons. Almost all GDP involve a ground ball in which the runner on first is put out, followed by a relay throw to first before the batter can arrive there. Stolen base attempts should decrease GDP, as any attempted steal of second (with less than two out) removes a likely GDP situation. Additionally, fast baserunners should be less likely, other things being equal, to hit into a GDP, because they arrive at first base more quickly than other runners.

Two sets of specifications are considered. The first set, numbered M1 through M4, contains four specifications for the production function, estimating coefficients on SB2 and CS2 separately for each model. The second set, numbered M1' through M4', uses specifications which estimates only coefficients for the number of attempts, where ATT2=SB2+CS2 and ATT3=SB3+CS3. For all models, each observation represents team-level statistics for one season, expressed on a per-game-played basis; all models have

5. `http://www.retrosheet.org`

6. In the 1974-1992 period, only seven teams collected at least 40 successful steals of third in a season. Leading the way is the 1976 Oakland Athletics, with 57 SB3; this team also has the maximum number of total successful steals (341) in the sample. This club was built around speed, and featured several first-class basestealers. The next six teams are the 1983 Oakland Athletics (SB3=54), 1985 St. Louis Cardinals (SB3=49), 1982 Oakland Athletics (SB3=45), 1986 St. Louis Cardinals (SB3=43), 1980 Montreal Expos (SB3=42), and 1988 St. Louis Cardinals (SB3=40). Each of these teams featured a player among the most prolific in stealing bases in history: the Athletics had Rickey Henderson, the Cardinals Vince Coleman, and the Expos Tim Raines. The teams with SB3 greater than 30 continue this pattern, with several more entries from clubs with these players present.

an intercept term (not reported).[7]

Table 3 provides descriptive statistics for all statistics in the regressions. Included are correlation coefficients of all statistics with triples, successful steals of second, and times caught stealing second, which are the three variables the previous analysis suggests were carrying information about team speed. Each of the newly-added explanatory variables have the expected correlations to 3B, SB2, and CS2.

The full model, Model 4, restores coefficient estimates on SB2 and CS2 to values which are consistent with equilibrium. The implied success percentage for a stolen base attempt to have zero net value is $\frac{.225}{.314} = .717$; this is comparable to the empirical success percentage of .655 overall during the sample period. Furthermore, the coefficient on CS2 is significantly different from zero in this model. In parallel, the full model, Model 4', of the set of models using attempts as regressors obtains a coefficient estimate on ATT2 that is not significantly different from the theoretical prediction of zero. Proceeding from Model 1 to Model 4 (resp., Model 1' to Model 4'), the coefficient estimates for SB2 and CS2 (resp., ATT2) change in the direction of implied success percentages consistent with equilibrium conditions.

				Correlation with		
Statistic	Mean	SD	Min/Median/Max	3B	SB2	CS2
R	4.26	0.45	3.10/4.24/5.53	0.14	-0.01	-0.18
1B	6.30	0.36	5.42/6.28/7.37	0.24	-0.00	0.03
2B	1.50	0.17	1.05/1.49/2.01	0.20	-0.01	-0.15
3B	0.23	0.07	0.07/0.22/0.49	1.00	0.31	0.27
HR	0.78	0.21	0.29/0.76/1.39	-0.23	-0.25	-0.30
BB	3.25	0.38	2.41/3.23/4.36	-0.15	0.04	-0.09
HBP	0.19	0.06	0.06/0.18/0.41	-0.09	-0.02	-0.05
GDP	0.76	0.11	0.47/0.76/1.07	-0.20	-0.48	-0.35
SB2	0.67	0.24	0.16/0.65/1.76	0.31	1.00	0.62
CS2	0.33	0.09	0.11/0.31/0.67	0.27	0.62	1.00
ATT2	1.00	0.31	0.32/0.98/2.42	0.33	0.97	0.78
SB3	0.07	0.05	0.00/0.06/0.35	-0.01	0.58	0.27
CS3	0.03	0.02	0.00/0.03/0.12	0.05	0.40	0.32
ATT3	0.10	0.07	0.00/0.09/0.45	0.01	0.58	0.31
AOE	0.34	0.07	0.19/0.34/0.53	0.26	0.38	0.35

Table 3. Descriptive statistics for the full 1974-1992 sample. All values are denominated in per-game-played terms.

7. With the exception of the 1981 season, all teams in the sample played between 159 and 164 games. The 1981 season was interrupted due to a player's strike; these 26 teams played between 103 and 111 games. Removing the 1981 season from the dataset does not substantively affect the results. As the number of outs per game is essentially constant across teams, so incorporating outs made into the regression would have a significant impact only on the unreported intercept term.

Stat	M1	M2	M3	M4	M1'	M2'	M3'	M4'	Stat
1B	0.522 (.019)	0.548 (.021)	0.552 (.020)	0.524 (.021)	0.521 (.019)	0.548 (.020)	0.554 (.021)	0.527 (.021)	1B
2B	0.689 (.042)	0.699 (.042)	0.702 (.042)	0.722 (.041)	0.708 (.042)	0.717 (.042)	0.715 (.041)	0.737 (.040)	2B
3B	1.066 (.112)	1.011 (.112)	1.056 (.114)	0.974 (.112)	1.056 (.113)	0.998 (.113)	1.059 (.113)	0.981 (.112)	3B
HR	1.472 (.038)	1.462 (.037)	1.461 (.037)	1.482 (.036)	1.476 (.038)	1.466 (.038)	1.463 (.037)	1.485 (.037)	HR
BB	0.341 (.018)	0.353 (.018)	0.349 (.018)	0.325 (.018)	0.348 (.018)	0.360 (.018)	0.354 (.018)	0.332 (.018)	BB
HBP	0.586 (.119)	0.570 (.118)	0.537 (.118)	0.505 (.115)	0.581 (.120)	0.564 (.119)	0.528 (.118)	0.495 (.115)	HBP
GDP		-0.222 (.074)	-0.215 (.074)	-0.207 (.072)		-0.238 (.074)	-0.222 (.073)	-0.217 (.072)	GDP
SB2	0.206 (.035)	0.163 (.038)	0.105 (.044)	0.089 (.043)	0.128 (.023)	0.088 (.026)	0.039 (.030)	0.008 (.030)	ATT2
CS2	-0.141 (.094)	-0.157 (.094)	-0.151 (.095)	-0.225 (.094)					
SB3			0.296 (.174)	0.216 (.180)			0.398 (.120)	0.338 (.118)	ATT3
CS3			0.492 (.415)	0.432 (.404)					
AOE				0.578 (.111)				0.551 (.111)	AOE
R^2	0.901	0.903	0.905	0.910	0.900	0.902	0.904	0.909	R^2
Adj. R^2	0.900	0.901	0.903	0.908	0.898	0.900	0.902	0.907	Adj. R^2
S. E.	0.143	0.142	0.141	0.137	0.144	0.143	0.141	0.138	S. E.
N	488	488	488	488	488	488	488	488	N

Table 4. Parameter estimates for four models of run scoring with various controls for team speed, using data from 1974-1992. The models are ordered in terms of increasing control. Standard errors for coefficient estimates are in parentheses.

The coefficients on SB3 and CS3 (resp., ATT3) should, in principle, should share a similar economic interpreation with those on SB2 and CS2 (resp., ATT2). However, the CS3 coefficient is positive and greater in magnitude than the coefficient on SB3, although the standard error on CS3 is sufficiently large that the estimate cannot be viewed as being significantly different from zero. The estimate for ATT3, however, is significant.

As noted earlier, the team-seasons in which large numbers of attempts to steal third are observed are not distributed randomly throughout the population, but are concentrated on time periods where teams featured one or more elite basestealers. While teams are not exactly identical from season to season, it is plausible that there is some other unobserved effect common to some or all of these teams. For now, this paper is satisfied with interpreting SB3 and CS3 (resp., ATT3) as variables which control for the presence of these unusual players.

4 Discussion

The introduction of additional production factors in the specification of the production function of a baseball team's offense removes a bias in the estimation of the value of a marginal stolen base attempt, and restores parameter estimates which are consistent with equilibrium.

The parameter estimates of Model 4 imply a success percentage which is slightly higher than the observed success percentage. Implicitly the analysis has assumed that the objective of the stolen base attempt is to maximize (or, for the defense, minimize) the expected number of runs scored, rather than the probability of winning the game. The distribution of future runs in an inning conditional on attempting a stolen base differs from that when a stolen base is not attempted; other things being equal, the distribution resulting from a stolen base attempt places more weight on low numbers of runs being scored. This is because suffering an out results in a large reduction of the probability of a big inning (many runs scored).[8]

In the early innings of a game, maximizing chances of winning is well-approximated by maximizing expected runs in the current inning. In the late innings, maximizing the probability of scoring becomes more relevant: in the extreme case of the bottom of the ninth inning of a tie game, the home team's objective reduces to maximizing the probability of scoring. Lindsey presents implied success percentages under an equilibrium assumption where the objective is to maximize probability of scoring: these success percentages are in the neighborhood of .55, which is lower than the corresponding probabilities under expected runs maximization.

While recovery of the equilibrium prediction for the marginal value of the stolen base attempt is reassuring for theory, this method suffers from the weakness that it cannot be used to evaluate the contribution of the stolen base strategy to a team's offense. For example, Berry uses the SB and CS coefficients he obtains to assign a value to a player's basestealing activity. However, what Berry measures is not the contribution of the player's stolen base activity, but rather an estimate of the player's additional contribution to his team due to speed-related factors *outside* basestealing. Thorn and Palmer apply a similar interpretation in using their SB and CS coefficients as part of a program to evaluate players based on their estimated marginal contributions. Since those authors use the values generated by Palmer's simulation, by definition the expected marginal contribution of a stolen base attempt is approximately zero, and much of the variation in a player's "stolen base runs" comes from the randomness presumably inherent in mixed strategy equilibrium.[9]

8. Lindsey's Table V provides distributions of this sort for his data.

Relatedly, KATSUNORI [6] extends a Markov model of player evaluation due to COVER AND KEILERS [5] to incorporate stolen bases and times caught stealing. Katsunori finds that for almost all players in the Japanese leagues, this addition actually reduces their estimated run production. The results in this paper help explain Katsunori's result. Presumably, the defense must pay some "enforcement cost" in order to deter stolen base attempts. This enforcement cost would reveal itself in the form of improved performance by the batter at the plate. Empirically, this is true; batters perform better with a runner on first base than with no runners on base. At the team level, the costs and benefits in this interaction are included; at the individual level, however, the benefits of an elite basestealer accrue at least in part to the *subsequent* hitter(s).

The methods of this paper, therefore, remain silent on the value of the ability to steal bases, because they are unable to identify how optimal strategies change as the characteristics of the players involved change. While this question is not particularly relevant for Walker and Wooders, as they take the players as given, it is of substantive interest to professional baseball teams, who face the problem of evaluating the potential contributions of players with very different skill sets towards the team's overall success. Understanding the stolen base in this context requires an adequate strategic model, which is pursued in the companion paper TUROCY [10].

Acknowledgments

The author wishes to thank Scott Berry, Donald Deere, Paula Hernandez-Verme, Steve Puller, John Straub, and Manuelita Ureta for valuable comments and encouragement in the development of this paper. The raw play-by-play data used in the paper is available for free on the World Wide Web at the Retrosheet website, `http://www.retrosheet.org`; the author gratefully acknowledges Retrosheet President Dave Smith and all Retrosheet volunteers for their efforts in producing and distributing this dataset. The scripts used to generate the team-level data, as well as a text file with these generated data, are available at the author's website, `http://econweb.tamu.edu/turocy`. All errors and omissions in this paper are the sole responsibility of the author.

Bibliography

[1] Jim Albert and Jay Bennett. *Curve Ball: Baseball, Statistics, and the Role of Chance in the Game*. Springer-Verlag, New York, 2001.

[2] Jay M. Bennett and John A. Flueck. An evaluation of Major League Baseball offensive performance models. *The American Statistician*, 37:76–82, 1983.

9. Note that a typical player might attempt around 10 stolen bases in a season, so there is a large sample standard deviation on a player's season success percentage.

[3] Scott M. Berry. Modeling Offensive Ability in Baseball. *Chance*, 13(4):56–59, 2000.

[4] 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.

[5] T. M. Cover and C. W. Keilers. An offensive earned-run average for baseball. *Operations Research*, 25:729–740, 1977.

[6] Ano Katsunori. Modified offensive earned-run average with steal effect for baseball. *Applied Mathematics and Computation*, 120:279–288, 2001.

[7] G. P. Lindsey. A scientific approach to strategy in baseball. In S. P. Ladany and R. E. Machol, editors, *Optimal Strategies in Sports*, pages 1–30. Elsevier-North Holland, New York, 1977.

[8] Charles M. Mottley. The Application of Operations-Research Methods to Athletic Games. *Journal of the Operations Research Society of America*, 2:335–338.

[9] John Thorn and Pete Palmer. *The Hidden Game of Baseball*. Doubleday & Company, Inc., Garden City NY, 1984.

[10] T. L. Turocy. The theory of theft: A strategic role for the stolen base in baseball. Mimeo, Texas A&M University, July 2003.

[11] M. Walker and J. Wooders. Minimax play at wimbledon. *American Economic Review*, 91:1521–1538, 2001.