Applied Economics, 2012, **44**, 707–716

Information effects in major league baseball betting markets

Matt E. Ryan[a,*], Marshall Gramm[b] and Nicholas McKinney[b]

[a]*Department of Economics and Quantitative Sciences, Duquesne University, 600 Forbes Avenue, Pittsburgh 15282, USA*
[b]*Department of Economics and Business, Rhodes College, Memphis, USA*

Previous studies point to a generally efficient baseball betting market with no profitable betting strategies. However, failure to consider the time of year in which the bets are placed neglects differences in available information throughout the season. This analysis largely confirms the general efficiency of the major league baseball betting market by existing measures; however, incorporating the time of the year in which the bet is made generates persistent profitable betting strategies. The process by which information impacts returns is considered; increasing difficulties in determining the true favourite likely play the largest role, while assessing the exact favourite underdog relationship also has an impact.

I. Introduction

While not covered to the extent of racetrack betting, a number of analyses characterize the nature of the betting market for baseball. In the first comprehensive analysis of legal gambling markets on major league baseball, Woodland and Woodland (1994) utilize money line odds to show that a reverse Favourite-Longshot (FL) bias exists, though the market as a whole presents no profitable wagering opportunity. The FL bias – typically used to describe racetrack betting markets – exists when the expected return to betting on favuorites exceeds that of betting on underdogs, or longshots. (See Sauer (1998) and Thaler and Ziemba (1988) for an overview of the FL bias literature.) Since Woodland and Woodland (1994) find a reverse FL bias, returns to betting on underdogs are greater than returns to betting on favourites. Updating their previous study, Woodland and Woodland (2003) further confirm the existence of the reverse FL bias absent a profitable betting strategy. Gandar *et al*. (2002) revisit Woodland and Woodland (1994) and provide evidence for the existence of a traditional FL bias in baseball wagering. Paul and Weinbach (2008), in examining line movements from opening to closing, find existence of a reverse FL bias amongst the uninformed betting public, and that informed bettors utilize this distortion to secure positive returns. Paul and Weinbach (2008), however, find this result to be true only in onshore markets (specifically, the Stardust sportsbook in Las Vegas) and not in offshore markets (specifically, Pinnacle sportsbook). Brown and Abraham (2002, 2004), Paul and Weinbach (2004), and Gandar and Zuber (2004), debate the nature of the over–under betting market in major league baseball.

With the exception of Paul and Weinbach (2008), the existing literature on baseball betting markets neglects to consider the role of evolving information on market efficiency and betting strategies. Moreover, Paul and Weinbach (2008) consider only changes in information (*via* early and late line movements) within the context of one particular game.

*Corresponding author. E-mail: ryanm5@duq.edu

Applied Economics ISSN 0003–6846 print/ISSN 1466–4283 online © 2012 Taylor & Francis
http://www.informaworld.com
DOI: 10.1080/00036846.2010.518951

To the knowledge of the authors, no existing study factors in the time when the bet is made with respect to the amount of the season that has elapsed.

The time of year when a bet is placed is crucial to consider. While market efficiency implies that bettors incorporate all available information, resulting in appropriate line movements and the erosion of any persistent profit opportunities, it does *not* say that the amount of information to be incorporated remains constant. Insofar as more information about the ability of a player or team is revealed as more games pass, the amount of the season that has elapsed necessarily plays a role on betting markets through its impact on the amount to be incorporated into betting decisions.

This scenario creates a particularly interesting outcome in baseball betting markets. In most sports betting markets – most notably basketball and football – bets are placed upon a team winning a game against a point spread. Distortions in information can cause movements in this point spread but will not change the return to any individual bet. Baseball betting markets, however, utilize a money line system for individual games. (For a description of the money line system, please see below.) Instead of handicapping games according to a point (or run) spread, bettors identify a winner per a sliding scale of odds. A team identified as an underdog will pay a higher return on a win than a favourite; the same is not true of a bet against a point spread. Therefore, variations in the amount of available information in baseball betting markets can distort the actual return on a bet, not simply the terms under which a constant return is earned.

Money line wagering

Betting markets on baseball operate through a system known as the money line. Instead of placing bets on the winner of a game in light of a point spread, as is common in basketball or football, bettors place wagers based on odds given to each team. Whereby betting markets based on point spreads pay equal returns to a winning underdog or favourite bet, money line wagers (like pari-mutuel wagers) pay more to winning underdogs than to winning favourites. Unlike the pari-mutuel wager, however, the bettor knows the return of a winning money line wager at the time he places the bet. Late line movements do not change the payout of an early money line bet.

Consider a sample money line: $(-160, +150)$. The negative number denotes the favourite and the positive number denotes the underdog. A winning bet of 160 units on the favourite will return 100 units; a winning bet of 100 units on the underdog will return 150 units. On a per unit basis, the winning favourite bet returns 0.625 units per unit bet, and the winning underdog bet returns 1.5 units per unit bet. Generalizing from the previous example, the money line $(-\beta_1, +\beta_2)$ pays $1/\beta_1$ to a winning favourite bet and β_2 to a winning underdog bet. The spread between the two lines – $|\beta_1| > |\beta_2|$ – generates commission for the book.

When considering the unique nature of the baseball betting market along with the nature of the game of baseball itself, incomplete information early in the season would seem to favour a strategy of betting on underdogs. Consider a sizeable shock to a money line: a favourite moves from -150 to -200. On a unit wager, the return to a winning bet falls from 0.67 to 0.5, or a difference of seventeen cents per dollar bet. However, consider the same money line movement for an underdog, from $+150$ to $+200$. On a unit wager, the return to a winning wager rises from 1.50 to 2.00, or a difference of 50 cents per dollar bet. Identical money line adjustments will have a larger impact upon the returns to underdog wagers.

Furthermore, given the increasing returns due to the money line betting system, betting on underdogs would seem to be a very profitable strategy. Using the same money lines from the example above, a win rate of 40% is needed for a money line of $+150$, and 33% for $+200$. In baseball, however, even the worst teams still win with a relatively high frequency. From 1999 to 2009, only 25 times did a team win less than 40% of its games – only 7.58% of the team-seasons.[1] Over the same period, only twice did a team win less than 33% of its games, or 0.61% of the team-seasons. However, of the available underdog bets over the same time span, 33.84% are at $+150$ or higher, and 8.66% are at $+200$ or higher.

Such aggregated analysis by no means proves a profitable betting strategy in targeting underdogs. Games with lines paying high returns for an underdog win are certainly not evenly distributed among all teams; it is only the teams believed to have an especially low chance of winning a particular game that receive such treatment. In addition, a team's season-long winning percentage does not imply a constant winning percentage against all opponents; any team's probability of winning a game declines as the proficiency of its opponent rises. Further, Woodland and Woodland (1994, 2003) find greater than expected losses in betting solely on underdogs, and Paul and Weinbach (2008) find no profitable

[1] Information concerning team performance comes from www.baseball-reference.com, the online Baseball Encyclopedia.

underdog betting strategy with offshore baseball markets in 2006.

However, difficulties arise when setting lines early in the season due to the shortage of information concerning team ability. Teams anticipated to be particularly successful could end up performing poorly, or vice versa. From 1999 to 2009, over 20% of the teams playing in the World Series had a losing record in the previous season. Should early season lines incorporate information from previous seasons to fill the void in current-season information, a blanket strategy as described above may prove effective. Given a large degree of uncertainty of the true underlying ability of professional baseball teams at the outset of the season, focusing on higher-paying returns – that is, underdogs – could be profitable.

This article proceeds as follows. Section II describes the data. Section III presents tests of market efficiency. Section IV analyses the returns to betting and considers these returns in light of the time of year when bets are placed. Section V concludes.

II. Data

The data for this analysis come from the Web site www.wagerline.com, an established offshore sportsbook. Paul and Weinbach (2008) utilize the same data source, albeit only for one season. Our data consist of 25 375 regular season major league baseball games from 1999 through 2009. Following from Woodland and Woodland (1994), Gandar *et al.* (2002) and Woodland and Woodland (2003), we analyse only money lines where $\beta_2 \geq +100$. This exclusion eliminates roughly 5% of the games played over this period. However, as our paper modifies the frameworks of these previous studies, this adjustment allows for the most direct comparison of our work with previous analyses.[2]

III. Tests of Market Efficiency

Within lines

If the baseball betting market is efficient, then the subjective probability (ρ) of winning a wager, or the probability of winning implied by the money line, at a particular line will not be statistically different from the objective probability (π) of winning a wager, or the percentage of wagers at that line that actually produced winning bets. In order to determine the efficiency of the market, we test the null hypothesis of $\rho_j = \pi_j$, where ρ_j is the subjective probability of winning at the jth line, and π_j is the objective probability of winning at the jth line. The subjective probability of an underdog win, from Gandar *et al.* (2002), is

$$\rho = (1 + \beta_1)/(2\beta_1 + \beta_1\beta_2 + 1) \tag{1}$$

where β_1 is the favourite line and β_2 is the underdog line. In order to ensure a normal distribution, only lines that satisfy $n_j\rho_j > 5$ and $n_j(1-\rho_j) > 5$ are analysed, where n_j is the number of games at the jth line. Lines that do not satisfy these requirements are excluded from the analysis. The appropriate test statistic is as follows:

$$z_j = (\pi_j - \rho_j) \Big/ \sqrt{\frac{\rho_j(1-\rho_j)}{n}} \tag{2}$$

The Appendix contains the results of this statistical test, and Table 1 compares these results to Woodland and Woodland (1994, 2003). Across the 165 unique lines that satisfy the restrictions to assume a normal distribution, 14 of them exhibit inefficiency at the 10% level. In comparing the percentage of lines found to be inefficient, this study finds a smaller percentage of inefficient lines than the original Woodland and Woodland study, but a higher percentage than the subsequent Woodland and Woodland analysis. Our study has the highest percentage of lines deemed inefficient at the 5% level – 4.85% of the lines in our study as compared to 3.85% in the original Woodland and Woodland article. Woodland and Woodland (2003) clearly shows the least amount of inefficiency of the three analyses.

As a whole, it would be difficult to suggest that the results of this study imply a large degree of market inefficiency. Woodland and Woodland (1994) note in their analysis that '[t]he market is remarkably efficient. Only 3 of the 26 lines tests lead to the rejection of the null hypothesis at the 10% level of significance.' If the 10% level of significance is the desired threshold, our study has less inefficiency in percentage terms. If the 5% level is appropriate, our study shows slightly more inefficiency. Ultimately, it would be difficult to argue that the offshore betting market analysed here exhibits a substantial difference in inefficiency from the Woodland and Woodland (1994) study using this particular test.

[2] A very small number of lines – less than one tenth of 1% – created illogical data points during the data collection process and were dropped from the set.

Table 1. Summary of within lines efficiency tests

	Appendix	Woodland and Woodland (2003)	Woodland and Woodland (1994)
	26 906	20 818	23 824
Total number of lines	165	31	26
Total number of inefficient lines	14	1	3
Percentage of inefficient lines (10%)	8.48%	3.23%	11.54%
Percentage of games in inefficient line categories (10%)	8.75%	5.34%	21.66%
Percentage of inefficient lines (5%)	4.85%	0.00%	3.85%
Percentage of games in inefficient line categories (5%)	7.21%	0.00%	6.01%

Note: Data for Woodland and Woodland (2003) show only new data not covered in Woodland and Woodland (1994).

Table 2. Across-lines bias test

α_0	−0.0151 (0.0194)
α_1	1.0368 (0.0438)
R^2	0.7750
F	1.90
p	0.1534

Note: Results from weighted (by n_j) least squares. The F-statistic, with (2, 163) degrees of freedom, from joint test of $\alpha_0 = 0$ and $\alpha_1 = 1$.

Across lines

To analyse market efficiency simultaneously across all lines we estimate the following model:

$$\pi_j = \alpha_0 + \alpha_1 \rho_j + \varepsilon_j \tag{3}$$

Estimating this equation in an ordinary least squares framework generates a significant amount of heteroskedasticity; the White test on the Ordinary Least Squares (OLS) estimation returns a χ^2 of 20.52, rejecting the null hypothesis of homoskedasticity at the 99% level. As such, we estimate the above regression equation using a weighted least squares procedure, whereby residuals from the jth line are weighted by the number of games at the jth line, n_j. Table 2 shows the results of this analysis. The null hypothesis for the test of market efficiency is the joint test of $\alpha_0 = 0$ and $\alpha_1 = 1$. The F-statistic from this joint test is 1.90 and at (2, 163) degrees of freedom, $p = 0.1534$; therefore the null hypothesis of market efficiency cannot be rejected at any reasonable level of significance. This result confirms the result of general market efficiency from the Section 'Model'. While neither α_0 nor α_1 are independent significantly, a reverse FL bias is suggested since $\alpha_0 < 0$ and $\alpha_1 > 1$. At higher values, objective probabilities exceed subjective probabilities, implying that favourites are overbet.

IV. Returns to Betting and the Time of Year

Model

The implicit assumption in previous studies is that the time of year that the bet is placed does play a role in outcomes; that is, the relationship between π and ρ is independent of the point in the season that the game takes place. However, should the time of year have an impact, then it should be reflected in varying the returns from bets over the course of the year. To isolate the impact of the time of year, we estimate the following model:

$$R_{tkj} = b_0 + b_{1tkj} + b_2 \Delta_k + b_3 Away_k + \varepsilon_{tkj} \tag{4}$$

This model follows from Gandar *et al.* (2002) in isolating the return on individual wagers. R is the actual return to a unit bet on the tth team in the kth game at the jth line. When the money line is −101 or less, the return to a winning bet is $1/\beta_j$ units. When the money line is +100 or greater, the return to a winning bet is β_j units. Any loss has a return of −1 units. ρ_{tkj} is the subjective probability of team t winning the game k at the jth line. Δ is the matrix of variables which capture the time during the season in which the kth game takes place. These measures include dummy variables to capture the month in which the kth game was played, as well as two additional dummy variables to capture games played in the first 2 weeks of the season and the last 2 weeks of the season.[3] *Away* is a dummy variable that measures whether the underdog in the kth game is the

[3] Baseball seasons typically start around 1 April and typically end at the end of September. Any games that occur during the month of March are included with the month of April; these games comprise 0.12% of the dataset. Any regular season games that occur in October (no postseason games are included in this dataset) are included with the month of September; these games comprise 1.14% of the dataset.

road team. According to Gandar *et al.* (2002), there is evidence that bettors treat home underdogs differently than road underdogs, and therefore should be controlled for in a regression framework. ε_{tkj} is the error term. The data are left censored at -1; the most a bettor can lose is 1 unit, while no theoretical upper bound exists. Therefore, a Tobit estimation is appropriate.

Expected values for estimates

Due to commission, every bet has an expectation of a negative return. Therefore, the expected value of b_0 is negative. The expected value of b_1 is a bit more subtle. Gandar *et al.* (2002) notes that the expected value of b_1 should be zero, and adds that a significant positive (negative) value is evidence of the traditional (reverse) FL bias. This reasoning is incorrect. (This reasoning would be correct if using a traditional OLS estimation; however, as Gandar *et al.* (2002, p. 1316) note: 'Since the regressand in this model is left-censored at -1, Tobit estimation is appropriate.') While bias will play a role in the exact final value of the b_1, its positive value is more a reflection of the reality that teams with higher subjective probabilities end up winning a higher percentage of games than teams with lower subjective probabilities. In other words, favourites are more likely to win than underdogs. As expected, favourites won 58.3% of games in our dataset. By winning more frequently, favourites generate more instances of positive returns as opposed to instances of losing 1 unit. Therefore, the expected value of b_1 is positive, not zero. Providing support to this claim is the fact that, of the 16 regressions estimating Equation 4 in both this analysis and Gandar *et al.* (2002), all 16 generate positive and extremely significant ($p<0.001$) estimates for b_1.

A simulation further confirms this expected value for b_1. Biases, like the FL bias, measure systematic discrepancies in subjective and objective probabilities. Consider the following scenario in which a bettor participates in a commission-free market. The bettor places four bets on a +300 underdog each time. The subjective probability of winning for a +300 underdog is 0.25. Assume now an objective probability equal to the subjective probability; therefore, $\pi=\rho=0.25$. One of the four bets wins for a return of 3 units, and the remaining three bets lose, for a return of -1 units per bet. The net return on these four bets is exactly zero. The bettor then places four more bets, now on a +100 team each time. The subjective probability of winning for a +100 team is 0.5. Assume again an objective probability equal to the subjective probability; now, $\pi=\rho=0.5$. Two of these four bets win, generating a return of 1 unit each, and the remaining two bets lose, each generating a return of -1. Again, the net return on these four bets is exactly zero. In this simple scenario, no bias exists in the data; subjective and objective probabilities are assumed equal, and no persistent profitable strategy exists. However, regressing actual returns on subjective probability using these eight observations generates an estimate for b_1 of 3.925; repeating this scenario to generate a dataset comparable to Woodland and Woodland (1994) produces an estimate with a level of significance well beyond the 1% level of confidence. A dataset with no bias or persistent profit opportunity generates a (extremely) statistically significant positive estimate. Furthermore, this hypothetical market exhibits the same positive and extremely significant estimates on subjective probability if bets are placed on both favourites as well as underdogs. The marginal effects also increase from estimating only favourites to estimating only underdogs, as per the analysis in the Section 'Results'. Thus, increases in marginal effects when considering underdogs as compared to favourites also does not constitute bias within a betting market.

With perfect information, the time of year when the bet is placed should play no role in determining the return. Since the constant, b_0, captures this persistent negative expected return across all bets, all values within vector b_2 should not be significantly different from zero. Per our argument in Section I, any significant deviation from the expectation of $b_2=0$ is evidence of information shortcomings in baseball betting markets and would highlight an inefficiency.

Results

Table 3 presents estimates and marginal effects for the Tobit model presented in the Section 'Model'. Estimates for the impact of the subjective probability (ρ) and of the away variable largely match the estimates in Gandar *et al.* (2002). ρ has a positive and extremely significant coefficient across all specifications, and its marginal impact is higher when considering underdogs as compared to favourites. *Away* fails to be significant in any specification.

The regressions that consider the time of year in which the bet is placed, however, provide compelling evidence that there are information inconsistencies early in the season that distort the betting market. The strongest evidence comes from games that occur during the first 2 weeks of the season. Returns are positive and significant at the 1% level when betting on all underdogs during the first 2 weeks of the season (Table 3), and correspondingly returns are

Table 3. Observations for all teams specifications clustered by game

	Coeff.	Marginal effects	Coeff.	Marginal effects	Coeff.	Marginal effects	Coeff.	Marginal effects
All teams								
Subjective probability	1.5279*** (0.1283)	0.7640	1.5281*** (0.1283)	0.7641	1.5283*** (0.1284)	0.7642	1.5283*** (0.1283)	0.7642
Away underdog			−0.0032 (0.0035)	−0.0016	−0.0033 (0.0035)	−0.0017	−0.0036 (0.0035)	−0.0016
Early April							0.0186*** (0.0069)	0.0093
April					0.0058 (0.0060)	0.0029		
May					−0.0036 (0.0057)	−0.0018		
June					0.0056 (0.0059)	0.0028		
August					−0.0026 (0.0061)	−0.0013		
September					−0.0016 (0.0063)	−0.0008		
Log-likelihood	−70 822.11		−70 822.09		−70 822.01		−70 821.96	
F	141.93		76.28		22.61		52.69	
p	0.0000		0.0000		0.0000		0.0000	
N	50 750		50 750		50 750		50 750	
Underdogs only								
Subjective probability	2.4632*** (0.3366)	1.0339	2.3948*** (0.3443)	1.0052	2.3551*** (0.3466)	0.9886	2.3786*** (0.3443)	0.9984
Away underdog			−0.0362 (0.0385)	−0.0152	−0.0392 (0.0385)	−0.0164	−0.0374 (0.0379)	−0.0157
Early April							0.1807*** (0.0680)	0.0758
April					0.1036 (0.0631)	0.0435		

May					−0.0719	−0.0302		
					(0.0612)			
June					0.0553	0.0232		
					(0.0614)			
August					−0.0500	−0.0210		
					(0.00610)			
September					−0.0118	−0.0050		
					(0.0611)			
Log-likelihood	−34 488.00		−34 487.56		−34 482.01		−34 484.16	
F	54.12		55.00		66.10		61.79	
p	0.0000		0.0000		0.0000		0.0000	
N	25 375		25 375		25 375		25 375	
Favourites only								
Subjective probability	0.9331***	0.5414	0.8932***	0.5183	0.8733***	0.5067	0.8845***	0.5133
	(0.1686)		(0.1725)		(0.1737)		(0.1725)	
Away underdog			0.0216	0.0125	0.0232	0.0135	0.0221	0.0128
			(0.0199)		(0.0199)		(0.0199)	
Early April							−0.0906**	−0.0526
							(0.0363)	
April					−0.0593*	−0.0344		
					(0.0327)			
May					0.0393	0.0228		
					(0.0314)			
June					−0.0272	−0.0158		
					(0.0317)			
August					0.0284	0.0165		
					(0.0312)			
September					0.0065	0.0038		
					(0.0313)			
Log-likelihood	−34 644.06		−34 643.47		−34 637.08		−34 640.35	
χ^2	30.70		31.88		44.67		38.14	
p	0.0000		0.0000		0.0000		0.0000	
N	25 375		25 375		25 375		25 375	

Notes: Tobit regression results, dependent variable is actual return to betting on team *t* in the *k*th game. Observations for All Teams specifications clustered by game.
***, ** and * denote statistic significance at the 1, 5 and 10% levels, respectively.

Table 4. Underdog betting returns by time of year

	N	R_U	$E(R_U)$	Z_1	Z_2
May through September					
All underdogs	20 999	−0.0180	−0.0173	−0.0839	−2.2297
Heavy underdogs	4607	0.0139	−0.0158	1.4780	0.6924
April					
All underdogs	3182	0.0307	−0.0178	2.4489**	1.5512*
Heavy underdogs	372	−0.0100	−0.0155	0.0798	−0.1475
Early April					
All underdogs	1153	0.0706	−0.0181	2.7244***	2.1683**
Heavy underdogs	146	0.2817	−0.0147	2.7104***	2.5759***

Notes: Z_1 is the test statistic for H_0: $R_U = E(R_U)$. Z_2 is the test statistic for H_0: $R_U \leq 0$. Heavy underdogs defined as $\beta_2 > +160$.
***, ** and * denote significance at the 1, 5 and 10% levels, respectively.

negative and significant at the 5% level when betting on all favourites over the same time period (Table 3). This 'early April' effect impacts the estimate for the entire month of April; the returns to betting on favourites over the entire month of April are negative and significant at the 10% level (Table 3), and the returns to betting on underdogs over the entire month of April is positive with $p = 0.101$ (Table 3). No other month carries any significance in any specification. This 'early April' effect lends credence to theory that sports books have a difficult time aggregating the proper information in setting lines correctly. As described in Section I, slight imperfections in setting baseball money lines are likely to favour the underdog more than the favourite. Our results bear out this scenario.

$$E(R_U) = \rho\beta_2 + (1 - \rho)(-1) = \rho(\beta_2 + 1) - 1 \quad (5)$$

Table 4 shows the expected returns, $E(R_U)$, and actual returns, R_U, to betting strategies by month. The above regression analysis suggests that a strategy focusing on April and early April could be a profitable strategy. Indeed, focusing on all underdogs in April generates a statistically significant rate of return above the expected rate at the 5% level of confidence. Furthermore, the strategy is profitable at the 10% level of confidence, and the rate of return is roughly 3%. When focusing only on early April, betting on all underdogs generates a rate of return above the expected level at the 1% level of confidence, and is a profitable strategy at the 5% level of confidence with a rate of return of roughly 7%. Finally, focusing only on heavy underdogs in the first half of April yields both a level of return above the expected level and a profitable strategy, both at the 1% level of confidence, with the rate of return at approximately 28%.

V. Conclusion

This analysis serves two primary purposes. First, existing efficiency studies on baseball betting markets end with data through the 1999 major league baseball season. Our analysis extends the understanding of baseball betting markets through the 2009 season. While our data come from offshore betting markets, they largely confirm the previous trends of generally high levels of efficiency across all lines.

However, previous analyses neglect to consider the time of the season in which the bets were placed. In theory, the time of the season in which the bet is placed could play a very important role. Information about the nature of teams is decidedly less while the season is young; this shortage of information could lead to persistent biases that generate profitable betting strategies. Indeed, during the first month of the season, betting on all underdogs yielded a statistically significant profit of roughly 3%. When focusing just on the first half of the month of April – when information is especially scarce – betting on all underdogs yielded a 7% return, and focusing only on heavy underdogs in the first half of April generated a staggering 28% return.

Focusing on the time of the season when bets are placed could provide numerous fruitful research opportunities – though, as mentioned in Section I, information differences throughout a season will have different impact depending on the betting system. In money line markets, shortages of information lead to difficulties in assessing true underdogs and favourites. Given the increased return to a unit bet on an underdog, this favours an underdog betting strategy. Further research could determine the impacts and the reasoning applied to traditional point spread markets.

References

Brown, K. H. and Abraham, F. J. (2002) Testing market efficiency in the major league baseball over-under betting market, *Journal of Sports Economics*, **3**, 311–19.

Brown, K. H. and Abraham, F. J. (2004) Testing market efficiency in the major league baseball over-under betting market, *Journal of Sports Economics*, **5**, 96–9.

Gandar, J. M. and Zuber, R. A. (2004) An evaluation of the debate over testing market efficiency in the major league baseball over-under betting market, *Journal of Sports Economics*, **5**, 100–5.

Gandar, J. M., Zuber, R. A., Johnson, R. S. and Dare, W. (2002) Re-examining the betting market on major league baseball games: is there a reverse favourite-longshot bias?, *Applied Economics*, **34**, 1309–17.

Paul, R. J. and Weinbach, A. P. (2004) Testing market efficiency in the major league baseball over-under betting market, *Journal of Sports Economics*, **5**, 93–5.

Paul, R. J. and Weinbach, A. P. (2008) Line movements and market timing in the baseball gambling market, *Journal of Sports Economics*, **9**, 371–86.

Sauer, R. D. (1998) The economics of wagering markets, *Journal of Economic Literature*, **36**, 2021–64.

Thaler, R. H. and Ziemba, W. T. (1988) Anomalies: parimutuel betting markets: racetracks and lotteries, *Journal of Economic Perspectives*, **2**, 161–74.

Woodland, L. M. and Woodland, B. M. (1994) Market efficiency and the favorite-longshot bias: the baseball betting market, *Journal of Finance*, **49**, 269–79.

Woodland, L. M. and Woodland, B. M. (2003) The reverse favourite-longshot bias and market efficiency in major league baseball: an update, *Bulletin of Economic Research*, **55**, 113–23.

Appendix

Within lines bias test

(β_{1j}, β_{2j})	ρ_j	n_j	w_j	π_j	z_j	(β_{1j}, β_{2j})	ρ_j	n_j	w_j	π_j	z_j
(−107, +100)	0.4917	130	63	0.4846	−0.1613	(−138, +130)	0.4285	131	62	0.4733	1.0353
(−108, +100)	0.4906	223	111	0.4978	0.2148	(−140, +130)	0.4270	987	417	0.4225	−0.2892
(−110, +100)	0.4884	1089	529	0.4858	−0.1720	(−139, +131)	0.4267	115	53	0.4609	0.7405
(−109, +101)	0.4882	185	92	0.4973	0.2471	(−140, +132)	0.4249	150	69	0.4600	0.8689
(−110, +102)	0.4859	190	94	0.4947	0.2441	(−141, +133)	0.4232	116	38	0.3276	−2.0834**
(−110, +103)	0.4847	131	64	0.4885	0.0892	(−143, +133)	0.4217	103	42	0.4078	−0.2871
(−111, +103)	0.4836	168	80	0.4762	−0.1916	(−142, +134)	0.4214	121	62	0.5124	2.0272**
(−112, +104)	0.4813	169	86	0.5089	0.7177	(−143, +135)	0.4197	113	42	0.3717	−1.0333
(−113, +105)	0.4790	274	131	0.4781	−0.0307	(−145, +135)	0.4183	810	367	0.4531	2.0090**
(−115, +105)	0.4770	1068	506	0.4738	−0.2094	(−144, +136)	0.4179	106	47	0.4434	0.5317
(−114, +106)	0.4768	175	98	0.5600	2.2040**	(−145, +137)	0.4162	105	53	0.5048	1.8409*
(−115, +107)	0.4746	170	81	0.4765	0.0499	(−147, +137)	0.4149	77	39	0.5065	1.6321
(−116, +108)	0.4724	167	84	0.5030	0.7930	(−146, +138)	0.4145	106	42	0.3962	−0.3820
(−117, +109)	0.4702	258	117	0.4535	−0.5370	(−147, +139)	0.4128	75	29	0.3867	−0.4600
(−118, +110)	0.4680	160	74	0.4625	−0.1397	(−148, +140)	0.4111	94	49	0.5213	2.1702**
(−120, +110)	0.4661	1067	488	0.4574	−0.5726	(−150, +139)	0.4108	69	33	0.4783	1.1382
(−119, +111)	0.4659	136	67	0.4926	0.6260	(−150, +140)	0.4098	513	208	0.4055	−0.2016
(−120, +112)	0.4637	255	115	0.4510	−0.4087	(−149, +141)	0.4095	99	36	0.3636	−0.9276
(−121, +113)	0.4616	159	78	0.4906	0.7317	(−150, +142)	0.4078	105	48	0.4571	1.0282
(−122, +114)	0.4596	129	59	0.4574	−0.0498	(−155, +140)	0.4067	261	108	0.4138	0.2333
(−123, +114)	0.4586	113	55	0.4867	0.5992	(−151, +143)	0.4062	68	30	0.4412	0.5874
(−123, +115)	0.4575	138	57	0.4130	−1.0479	(−153, +142)	0.4059	76	35	0.4605	0.9692
(−125, +115)	0.4557	894	410	0.4586	0.1751	(−152, +144)	0.4046	67	26	0.3881	−0.2754
(−124, +116)	0.4554	165	80	0.4848	0.7587	(−153, +145)	0.4030	82	25	0.3049	−1.8108*
(−125, +117)	0.4534	145	66	0.4552	0.0429	(−155, +145)	0.4017	574	233	0.4059	0.2048
(−126, +118)	0.4514	140	67	0.4786	0.6464	(−154, +146)	0.4014	95	37	0.3895	−0.2365
(−127, +118)	0.4505	109	42	0.3853	−1.3681	(−155, +147)	0.3998	72	31	0.4306	0.5331
(−127, +119)	0.4494	119	53	0.4454	−0.0880	(−157, +146)	0.3996	58	30	0.5172	1.8299*
(−128, +120)	0.4474	152	64	0.4211	−0.6535	(−160, +145)	0.3988	121	47	0.3884	−0.2324
(−130, +120)	0.4457	1120	495	0.4420	−0.2540	(−156, +148)	0.3982	71	31	0.4366	0.6612
(−129, +121)	0.4454	110	46	0.4182	−0.5754	(−157, +149)	0.3966	53	25	0.4717	1.1169
(−130, +122)	0.4435	182	71	0.3901	−1.4499	(−160, +148)	0.3959	80	30	0.3750	−0.3815
(−131, +123)	0.4416	121	54	0.4463	0.1043	(−158, +150)	0.3951	76	28	0.3684	−0.4757
(−133, +123)	0.4400	112	51	0.4554	0.3282	(−160, +150)	0.3939	541	218	0.4030	0.4293
(−132, +124)	0.4397	116	46	0.3966	−0.9354	(−159, +151)	0.3936	65	26	0.4000	0.1062

(*continued*)

Continued

(β_{1j}, β_{2j})	ρ_j	n_j	w_j	π_j	z_j	(β_{1j}, β_{2j})	ρ_j	n_j	w_j	π_j	z_j
(−133, +125)	0.4378	116	40	0.3448	−2.0176**	(−160, +152)	0.3920	85	35	0.4118	0.3725
(−135, +125)	0.4362	922	402	0.4360	−0.0114	(−163, +151)	0.3913	76	28	0.3684	−0.4086
(−134, +126)	0.4359	134	60	0.4478	0.2773	(−165, +150)	0.3911	127	49	0.3858	−0.1228
(−135, +127)	0.4340	129	43	0.3333	−2.3073**	(−161, +153)	0.3905	81	30	0.3704	−0.3718
(−137, +127)	0.4325	106	43	0.4057	−0.5576	(−162, +154)	0.3890	70	21	0.3000	−1.5278
(−136, +128)	0.4322	138	67	0.4855	1.2648	(−163, +155)	0.3875	61	28	0.4590	1.1460
(−137, +129)	0.4303	110	44	0.4000	−0.6426	(−165, +155)	0.3864	408	144	0.3529	−1.3895
(−164, +156)	0.3861	58	28	0.4828	1.5127	(−195, +180)	0.3508	30	9	0.3000	−0.5827
(−167, +155)	0.3854	56	19	0.3393	−0.7085	(−196, +183)	0.3480	24	7	0.2917	−0.5789
(−165, +157)	0.3846	56	16	0.2857	−1.5209	(−200 +185)	0.3448	132	44	0.3333	−0.2778
(−170, +155)	0.3838	100	42	0.4200	0.7445	(−201, +185)	0.3445	95	28	0.2947	−1.0197
(−166, +158)	0.3831	55	21	0.3818	−0.0200	(−202, +185)	0.3441	43	13	0.3023	−0.5764
(−167, +159)	0.3817	55	23	0.4182	0.5572	(−200, +188)	0.3425	225	93	0.4133	2.2401**
(−170, +158)	0.3810	63	22	0.3492	−0.5202	(−206, +188)	0.3403	41	15	0.3659	0.3457
(−168, +160)	0.3802	65	28	0.4308	0.8390	(−205, +189)	0.3399	15	5	0.3333	−0.0533
(−170, +160)	0.3792	349	140	0.4011	0.8445	(−205, +190)	0.3391	44	15	0.3409	0.0257
(−169, +161)	0.3788	53	21	0.3962	0.2612	(−208, +190)	0.3380	25	8	0.3200	−0.1904
(−170, +162)	0.3774	65	24	0.3692	−0.1361	(−210, +190)	0.3373	130	49	0.3769	0.9550
(−175, +160)	0.3767	110	43	0.3909	0.3073	(−210, +192)	0.3358	36	13	0.3611	0.3217
(−171, +163)	0.3760	55	18	0.3273	−0.7462	(−210, +193)	0.3350	166	52	0.3133	−0.5943
(−175, +162)	0.3749	56	27	0.4821	1.6575*	(−210, +195)	0.3335	26	8	0.3077	−0.2793
(−172, +164)	0.3746	54	19	0.3519	−0.3456	(−215, +195)	0.3318	33	10	0.3030	−0.3515
(−173, +165)	0.3732	49	20	0.4082	0.5056	(−214, +196)	0.3314	18	7	0.3889	0.5180
(−175, +165)	0.3723	309	123	0.3981	0.9385	(−215, +197)	0.3303	29	9	0.3103	−0.2290
(−174, +166)	0.3719	48	18	0.3750	0.0450	(−215, +200)	0.3281	23	5	0.2174	−1.1310
(−175, +167)	0.3705	41	17	0.4146	0.5852	(−218, +200)	0.3272	26	6	0.2308	−1.0476
(−180, +165)	0.3699	67	23	0.3433	−0.4510	(−220, +200)	0.3265	249	79	0.3173	−0.3116
(−176, +168)	0.3691	52	22	0.4231	0.8060	(−220, +202)	0.3251	23	8	0.3478	0.2330
(−180, +167)	0.3681	49	19	0.3878	0.2848	(−220, +204)	0.3236	35	16	0.4571	1.6883*
(−177, +169)	0.3678	26	11	0.4231	0.5845	(−225, +205)	0.3214	16	3	0.1875	−1.1467
(−178, +170)	0.3665	43	13	0.3023	−0.8729	(−225, +207)	0.3200	28	11	0.3929	0.8269
(−180, +170)	0.3655	254	95	0.3740	0.2807	(−230, +210)	0.3164	205	74	0.3610	1.3725
(−179, +171)	0.3651	35	10	0.2857	−0.9759	(−230, +212)	0.3150	45	9	0.2000	−1.6608*
(−180, +172)	0.3638	53	16	0.3019	−0.9373	(−235, +215)	0.3116	26	8	0.3077	−0.0425
(−185, +170)	0.3633	61	18	0.2951	−1.1076	(−240, +215)	0.3102	49	11	0.2245	−1.2973
(−181, +173)	0.3625	21	8	0.3810	0.1757	(−235, +217)	0.3102	17	3	0.1765	−1.1920
(−185, +172)	0.3616	43	17	0.3953	0.4608	(−240, +220)	0.3069	135	40	0.2963	−0.2661
(−182, +174)	0.3612	30	14	0.4667	1.2023	(−240, +222)	0.3055	23	9	0.3913	0.8930
(−183, +175)	0.3599	26	13	0.5000	1.4879	(−250, +220)	0.3043	38	14	0.3684	0.8584
(−185, +175)	0.3591	209	81	0.3876	0.8590	(−250, +230)	0.2979	144	37	0.2569	−1.0739
(−184, +176)	0.3587	36	15	0.4167	0.7257	(−260, +230)	0.2956	22	10	0.4545	1.6342
(−185, +177)	0.3574	34	9	0.2647	−1.1277	(−260, +240)	0.2894	104	33	0.3173	0.6278
(−190, +175)	0.3569	88	32	0.3636	0.1314	(−270, +250)	0.2814	77	21	0.2727	−0.1686
(−186, +178)	0.3561	21	6	0.2857	−0.6739	(−280, +260)	0.2738	61	12	0.1967	−1.3497
(−190, +176)	0.3561	27	7	0.2593	−1.0508	(−300, +270)	0.2649	44	10	0.2273	−0.5656
(−190, +180)	0.3528	195	66	0.3385	−0.4190	(−300, +275)	0.2623	46	11	0.2391	−0.3572
(−191, +180)	0.3524	101	35	0.3465	−0.1231	(−320, +290)	0.2518	36	10	0.2778	0.3591
(−192, +180)	0.3520	24	9	0.3750	0.2362						

Notes: β_{1j} and β_{2j} are the *j*th favourite and underdog lines, while ρ_j, n_j, w_j and π_j are the subjective win probabilities, number of games, number of underdog wins and the objective probabilities at the *j*th line, respectively.
$\Sigma n_j = 24\,773$.
** and * denote significance at the 5 and 10% levels, respectively.