What Causes the Favorite-Longshot Bias? Further Evidence from Tennis

Lahvicka, Jiri

30 June 2013

What Causes the Favorite-Longshot Bias? Further Evidence from Tennis

Jiří Lahvička

E-mail: jirka@lahvicka.cz

In sports betting markets, bets on favorites tend to have a higher expected value than bets on longshots. This article uses a data set of almost 45,000 professional single tennis matches to show that the favorite-longshot bias is much stronger in matches between lower-ranked players, in later-round matches, and in high-profile tournaments. These results cannot be solely explained by bettors being locally risk-loving or overestimating chances of longshots, but are consistent with bookmakers protecting themselves against both better informed insiders and the general public exploiting new information.

Keywords: favorite-longshot bias, tennis, sports betting, market efficiency

JEL classification: L83; G14

Last revision: June 30th, 2013

1 Introduction

In sports betting markets, bets on favorites usually have a higher expected value than bets on longshots (Sauer, 1998; Cain et al., 2003). There are three types of explanations for this so-called favorite-longshot bias (Snowberg and Wolfers, 2010; Makropoulou and Markellos, 2011). The first explanation claims that bettors are local risk-lovers and bookmakers take advantage by lowering the odds on longshots. According to the second explanation, bettors overestimate winning probabilities of longshots and bookmakers again take advantage of this psychological bias. The third explanation is based on information asymmetry; bookmakers could potentially lose a lot of money if they underestimate longshots and this mispricing is exploited by either better informed insiders or by the general public reacting faster than bookmakers to new information. Therefore, bookmakers offer lower odds on longshots to protect themselves against this type of loss.

To distinguish between these competing explanations, this article uses a data set of almost 45,000 professional single tennis matches to show that the favorite-longshot bias is much more pronounced in matches between lower-ranked players, in later-round matches, and in high-profile tournaments. These results, as discussed later, are consistent with the information asymmetry explanation. The favorite-longshot bias in tennis was already analyzed by Forrest and McHale (2007), but they had a much smaller data set, did not test the effect of players' ranks or tournament round, and did not find any difference for high-profile tournaments.

2 Data

The data set consists of results of 44,871 professional men's and women's single tennis matches with valid betting odds.[1] The decimal betting odds on each player's win were converted to implied probabilities of winning in the standard way by calculating their inverse values. Since the two resulting numbers for each match add up to more than one to allow the bookmaker to have profit, they have to be both divided by their sum. Because the two possible bets on each match are not independent (implied probabilities add up to one, exactly one bet pays off), only one (chosen randomly) is included in the final data set. Therefore, there are 44,871 observed bets with an implied probability of the player winning (variable *ImpliedProbability*) and a corresponding match result (variable *Result* that equals one if the player won and zero if the player lost).

To test how the favorite-longshot bias differs across various types of matches, the following dummy variables are defined: *LowerRank* equals one in 12,878 matches where both players were outside of

[1] The data set was downloaded from the website tennis-data.co.uk on June 22nd, 2013. The men's tennis matches start in 2002; the women's matches start in 2007. The betting odds are the latest available odds by the bookmaker Bet365. Originally, there were 48,042 matches, but 3,171 matches (6.6%) were discarded due to missing odds or a withdrawal of one player before the match started.

top 50 in ATP/WTA rankings, zero otherwise; *LaterRound* equals one in 24,189 matches that were not in the first round (lowest round in the data set), zero otherwise; and *HighProfile* equals one in 8,962 matches in a high-profile tournament (Grand Slam, ATP World Tour Finals, or WTA Tour Championships), zero otherwise.

3 Model and Results

To test whether the favorite-longshot bias exists in the market as a whole, the following standard linear probability model is employed:

$Result = \beta_0 + \beta_1 * ImpliedProbability + \varepsilon$

In the absence of bias (null hypothesis), the coefficient values would be $\beta_0 = 0$ and $\beta_1 = 1$, while the standard favorite-longshot bias would be indicated by $\beta_0 < 0$ and $\beta_1 > 1$. The estimation results[2] in Table 1 show that the favorite-longshot bias is indeed present in the investigated data set; the winning probability implied by the betting odds is lower than the actual probability in case of longshots and higher than the actual probability in case of favorites.

	Coefficient	Standard Error
Constant	-0.0293***	0.0044
ImpliedProbability	1.0594***	0.0077

Table 1: The Favorite-Longshot Bias in the Whole Market, N = 44,871

To investigate whether the favorite-longshot bias differs across various types of matches, the model is expanded in the following way:

$Result = \beta_0 + \beta_1 * ImpliedProbability + \beta_2 * LowerRank + \beta_3 * LowerRank * ImpliedProbability + \beta_4 * LaterRound + \beta_5 * LaterRound * ImpliedProbability + \beta_6 * HighProfile + \beta_7 * HighProfile * ImpliedProbability + \varepsilon$

In case of no difference among various types of matches (the null hypothesis), $\beta_2 \ldots \beta_7 = 0$, while the overall bias would still be captured by $\beta_0 < 0$ and $\beta_1 > 1$. The estimation results for the expanded model are presented in Table 2.

[2] The estimation method in the whole article is OLS with heteroskedasticity-robust standard errors. One star indicates p-value < 0.1, two stars p-value < 0.05, three stars p-value < 0.01.

	Coefficient	Standard Error
Constant	0.0032	0.0085
ImpliedProbability	1.0051	0.0148
LowerRank	-0.0539***	0.0132
LowerRank * ImpliedProbability	0.0918***	0.0244
LaterRound	-0.0233**	0.0092
LaterRound * ImpliedProbability	0.0358**	0.0162
HighProfile	-0.0361***	0.0094
HighProfile * ImpliedProbability	0.0704***	0.0160

Table 2: The Favorite-Longshot Bias across Various Types of Matches, N = 44,871

The coefficients show the favorite-longshot bias is much stronger in matches between lower-ranked players, in later-round matches, and in matches in high-profile tournaments, while it is practically nonexistent in the other matches. These results are robust across different model specifications and have also been confirmed by comparing average implied probabilities with relative frequencies of winning over different probability ranges for different types of matches (similarly to Forrest and McHale 2007). A graphical analysis also confirms that the relationship between the implied and actual probability of a win is approximately linear.

4 Discussion

The results seem to be contradictory; on the one hand, the favorite-longshot bias is stronger in later-round matches and in matches in high-profile tournaments, i.e. in matches that are likely to attract high betting volumes; on the other hand, the favorite-longshot bias is also more pronounced in matches between lower-ranked players, which are likely to exhibit low betting volumes. This pattern cannot be explained solely by people being local risk-lovers or overestimating chances of longshots; if all bettors had the same preferences or biases, the type of match should not matter at all. Even if the risk-loving preferences (or the corresponding bias) were exhibited only by occasional bettors, thus causing the stronger favorite-longshot bias in matches that are likely to attract high betting volumes, it would not explain why the bias is also more pronounced in matches between lower-ranked players. Therefore, at last one part of the explanation must lie in the information asymmetry.

Forrest and McHale (2007) argued that in Grand Slam tournaments, players are more motivated and less likely to underperform, so the role of private information should be much smaller. Consequently, if the favorite-longshot bias was a defense of bookmakers against better informed insiders, it should be smaller in high-profile tournaments. However, according to the results in this article, the bias is actually larger. This is hard to explain as a defense against insider trading; besides players being more motivated, the proportion of insiders among all bettors is also likely to be smaller, not larger, in high-profile tournaments.

The most plausible explanation of the results seems to be a combination of two information asymmetry approaches: Matches between lower-ranked players are harder to predict, since public information is limited and private information about players' motivation or health problems could play a large role; therefore, it makes sense for the bookmaker to set lower odds on the longshot to minimize possible losses. On the other hand, private information should not play such a big role in later tournament rounds and high-profile tournaments, but in such matches the bookmaker faces a different kind of risk; the general public could react faster than the bookmaker to newly available information. Combined with a high volume of bets, this could mean a considerable loss, so the bookmaker again protects itself by setting lower odds on the longshot.

Of course, the information asymmetry explanation does not rule out that the other alternatives, i.e. risk-loving preferences or overestimating small probabilities of winning, also play a role. Clearly, more research is needed. One possible direction would be to test more thoroughly whether the stronger favorite-longshot bias in high-profile tournaments also exists in other individual or team sports (or even in tennis doubles); if the above explanation is correct, the effect in team sports should be smaller, since the impact of new information (e.g. a minor sickness of a player) is likely to have less influence on the expected result.

References

Cain, M., Law, D., & Peel, D. (2003). The Favourite-Longshot Bias, Bookmaker Margins and Insider Trading in a Variety of Betting Markets. *Bulletin of Economic Research, 55*(3), 263-273.

Forrest, D., & McHale, I. (2007). Anyone for Tennis (Betting)? *The European Journal of Finance, 13*(8), 751-768.

Makropoulou, V., & Markellos, R. N. (2011, September). Optimal Price Setting in Fixed-Odds Betting Markets under Information Uncertainty. *Scottish Journal of Political Economy, 58*(4), 519-536.

Sauer, R. D. (1998, December). The Economics of Wagering Markets. *Journal of Economic Literature, 36*(4), 2021-2064.

Snowberg, E., & Wolfers, J. (2010). Explaining the Favorite-Long Shot Bias: Is it Risk-Love or Misperceptions? *Journal of Political Economy, 118*(4), 723-746.