European Journal of Operational Research 249 (2016) 1102–1112

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

Estimating risk preferences of bettors with different bet sizes[☆]

Eberhard Feess[a,*], Helge Müller[a], Christoph Schumacher[b]

[a] Frankfurt School of Finance & Management, Frankfurt D-60314, Germany
[b] School of Economics and Finance, Massey University, Auckland, New Zealand

ARTICLE INFO

Article history:
Received 16 April 2014
Accepted 25 September 2015
Available online 5 October 2015

Keywords:
Applied probability
Betting markets
Favorite-longshot bias
Estimation of risk preferences
Overweighting of small probabilities

ABSTRACT

We extend the literature on risk preferences of a representative bettor by including odds-dependent bet sizes in our estimations. Accounting for different bet sizes largely reduces the standard errors of all coefficients. Substituting the coefficients from the model with equal bet sizes into the model with odds-dependent sizes leads to a sharp decline in the likelihood which shows that accounting for different amounts is important. Our estimations strongly reject the hypothesis that the overbetting of outcomes with low probabilities (favorite-longshot bias) can be explained by risk-seeking bettors. Depending on the exact specification within cumulative prospect theory, the data can best be described by an overweighting of small probabilities which is more pronounced in the gain domain. Models allowing for two parameters for probability weighting each in the gain- and in the loss domain are superior.

1. Introduction

Sports betting data is useful for estimating risk preferences due to the fact that the risk of different bets is uncorrelated, and because returns are realized soon after the bets are placed. Consequently, sports betting data has often been applied for estimating which of the canonical models for choices under risk (in particular expected utility theory, EUT, and cumulative prospect theory, CPT) fits the data best. A rather robust result is that different specifications of CPT perform better than EUT, a finding that coincides with most experimental results on behavior under risk (see Harrison & Rutström (2008), for a comprehensive overview).

Most of the literature adopts the so-called representative bettor approach which estimates the parameters of the bettors' utility functional under three assumptions: (i) odds are explained by the bettors' indifference condition between the available choices for an event, (ii) all bettors are identical, and (iii) bet sizes are independent of odds. Straightforwardly, assumptions (ii) and (iii) are counterfactual and are mainly adopted due to data limitations. A few recent studies drop assumption (ii) and estimate individual preferences either by using data on individual betting behavior (Andrikogiannopoulou & Papakonstantinou, 2013) or by estimating heterogeneity from aggregated data (Chiappori, Gandhi, Salanié, & Salanié, 2012). Our work is complementary to these papers as we keep the assumption of a representative bettor but drop assumption (iii): we estimate the representative bettors' utility functional from the indifference condition between choices which differ both in odds and bet sizes. We can do so as our data set includes the actual amount for each bet placed.

To the best of our knowledge, our paper is the first that estimates risk preferences within the representative bettor approach for CPT with different bet sizes. The descriptive statistics for our large data set that includes approximately 800,000 observations already reveals that accounting for different bet sizes is important. Bet sizes are decreasing in odds to such a large degree that the correlation between odds and the variance of return, which may be seen as a rough proxy for risk, is negative. Although this is not surprising as few (non-professional) bettors would be willing to bet large amounts on longshots, it is an important observation which implies that it cannot be taken for granted that those betting on longshots take higher risks than those betting on favorites.

We find that specifications including the actual bet sizes reduce the standard errors considerably. Furthermore, when we substitute the coefficients estimated from the model with equal bet sizes into the model with actual bet sizes, we find a strong decline in the likelihoods. This shows that accounting for different bet sizes is important.

In line with most of the data utilized in the literature (see the overview in Ottaviani & Sørensen (2008)),[1] we observe a

☆ We are grateful to an anonymous referee for very helpful comments. We also thank seminar participants at the annual meeting of the European Economic Association in Toulouse 2014, at WHU in Valendar and at Massey University in Auckland.

* Corresponding author. Tel.: +49 69154008398.
E-mail addresses: e.feess@fs.de, eberhard_feess@yahoo.de (E. Feess), h.mueller@fs.de (H. Müller), C.Schumacher@massey.ac.nz (C. Schumacher).

[1] Only few papers find a reversed FLB, i.e. markets where average returns are higher when betting on longshots (see in particular Woodland and Woodland (1994), and

http://dx.doi.org/10.1016/j.ejor.2015.09.053
0377-2217/

pronounced favorite-longshot bias (FLB) expressing that returns are, on average, decreasing in odds. When we define favorites as the bottom fifty percent lowest odds placed and longshots as the other fifty percent, losses with longshots are on average around 21 percent, compared to only 8 percent for favorites. When estimating the preferences based on EUT, the existence of a FLB necessarily leads to the result that bettors are risk-seeking. EUT, however, performs rather poorly both for equal and with actual bet sizes, i.e. it leads to likelihoods far above those for CPT. The CPT-estimations suggest risk-neutral bettors, so that EUT can contribute little to our understanding of the FLB. By contrast, models including overweighting of small probabilities, separated by the gain- and the loss domain, fit the data well.

Our findings depend to some degree on the exact specification of the probability weighting functions, but the following results can be seen as robust (see Section 4 for details): First, low probabilities are overweighted relative to high probabilities. This holds both in the gain- and in the loss domain, and is consistent with the existence of a pronounced FLB. Second, probabilities in the gain domain are overweighted relative to probabilities in the loss domain. This sheds (some) light on why people invest at all in risky assets with negative expected value. In line with this, specifications including loss aversion show that the representative bettor puts more weight on gains compared to losses. Third, the performance of models, captured by the likelihood, improves when we allow for two parameters for probability weighting each in the gain- and in the loss domain. Thereby, we need to take into account that models with more parameters automatically lead to a (weakly) higher likelihood. To correct for this effect, we apply the widely used Akaike Information Criterion (AIC), but we still find that more complex weighting-functions are superior for reasons we will discuss in Section 4.

Besides the preference-based approach, there are several other important explanations of the FLB, including those considering the profit-maximization of poorly- (Shin, 1991, 1992) or well-informed (Levitt, 2004) bookmakers, and those based on the heterogeneity in the beliefs of bettors with respect to the winning probabilities (Ali, 1977; Gandhi & Serrano-Padial, 2014; Ottaviani & Sørensen, 2009). In Section 5, we will briefly relate our findings to these approaches. To streamline the following literature review, we now restrict attention to the preference-based approaches.

The early literature estimated the parameters of the representative bettor's utility functional only for EUT (Ali, 1977; Weitzman, 1965). Jullien and Salanié (2000) have shown that models based on CPT outperform EUT (see the overviews in Jullien & Salanié, 2008; Ottaviani & Sørensen, 2008). In the part of their paper in which they replicate the Jullien and Salanié (2000)-approach, Gandhi and Serrano-Padial (2014) find no evidence for risk-seeking preferences, something that coincides with our results. Snowberg and Wolfers (2010) use exotic bets (for instance bets on the winner and the second best horse, called "exacta") for discriminating between explanations based on utility functions and on probability weighting. They also confirm that non-linear probability weighting explains the data best.

Bradley (2003) develops a theoretical model that allows for odds-dependent bet sizes, but has no data on those. With data from betfair.com, Kopriva (2009) also reports that bet sizes are largely decreasing in odds, but he estimates the representative bettor's preferences only with EUT but not with CPT. To our knowledge, there is no other paper that extends the comparison of EUT and CPT within the representative bettor approach to different bet sizes.

There are at least two caveats in assessing our results: with any other paper on betting, we first share the problem that bettors are unlikely to be representative for the population as a whole, so that there is a sample selection bias. In particular, while the overweighting of small compared to large probabilities is a well-known phenomenon, the overweighting of gains compared to losses might be specific for betting markets. Thus, our results indicate in which respects betting markets are similar and different to other choices under risk. Second, we interpret CPT as a statistical tool rather than as an explanation for the underlying cognitive processes of bettors. In other words, we find that bettors behave *as if they* overweight small probabilities, but our analysis keeps silent on whether this is attributable to systematic misperceptions of probabilities, to preferences for weak contestants, or to other reasons not accounted for in our estimations. We will get briefly back to this in Section 5. Here, we emphasize that, whenever we talk of e.g. the "overweighting of small probabilities", we use it in the cautious sense just mentioned.

The remainder of our paper is organized as follows: Section 2 describes the data set. Section 3 explains the methodology. Section 4 presents our results for the benchmark model, and discusses robustness checks. Section 5 compares our findings to other approaches. We conclude in Section 6.

2. Data

Our data set was compiled in close cooperation with the 'New Zealand Racing Board' (NZRB) which is the only licensed betting agency in New Zealand. The initial data contain all 5,136,660 fixed-odds bets placed at the agency between August 2006 and April 2009. By contrast to parimutuel betting where the total amount is shared among all successful bettors, odds in fixed-odds-betting are set by bookmakers. An advantage of fixed-odds-betting for our purposes is that bettors know the final odds at the time of betting. A disadvantage is that odds are not solely determined by the demand of bettors, but also by strategic considerations of bookmakers (see Direr, 2013; Levitt, 2004; Shin, 1991, 1992). Data with fixed-odds-betting is also used by Jullien and Salanié (2000), Kopriva (2009) and Andrikogiannopoulou and Papakonstantinou (2013).

For each event included in our analysis, we have information on odds, the number of bets on each possible outcome, the outcome itself and the bet sizes for each bet. From outcomes, we calculate success probabilities for odds. To include bet sizes in the representative bettor approach, we calculate the average bet sizes for the respective odds.

Many events in our data set have more than two possible outcomes, for instance games where draws are possible as in Soccer, or competitions with many contestants such as sailing and golf. As the estimation of risk preferences is based on the indifference condition for all outcomes, the processor time is disproportionately increasing in the number of possible outcomes.[2] Therefore, we restrict attention to events with a maximum of six outcomes.

When calculating the number of bets and average bet sizes per outcome, we furthermore excluded all bets where either the average bet size or the average potential gain exceeded 500 NZ dollar.[3] There are two reasons for this: first, we are interested in the risk preferences of the average casual bettor, and we do not want to mix up casual and professional bettors. Second, large single bets would otherwise dominate (and potentially distort) the results when using average bet sizes. We also ran estimations with many different thresholds and for a model including all bets. Results are qualitatively robust and several specifications are presented in the Appendix. We exclude events with less than five bets and events with irregularities that could not be clarified with the bookmaker.

Sobel and Raines (2003)). Boulier, Stekler, and Amundson (2006) find no significant impact of odds on returns.

[2] Each estimation took around thirty minutes, and the number of indifference equations with n outcomes is $\binom{n}{2}$.

[3] The exchange rate of the NZ dollar to the US dollar fluctuated over the observation period, but on average, 1NZ dollar was about 65 Cent.