Algorithmic Social Sciences Research Unit

ASSRU

Department of Economics
University of Trento
Via Inama 5
381 22 Trento Italy

DISCUSSION PAPER SERIES

7 – 2011/II

AGENT-BASED MODELLING OF THE PREDICTION MARKETS♠

Tongkui Yu & Shu-Heng Chen

JULY 2011

♠ Text of the talk given at the ASSRU/Department of Economics Seminar, University of Trento, 26 May, 2011. Professor Shu-Heng Chen is a *Founding Honorary Associate* of ASSRU.

Agent-Based Modeling of the Prediction Markets

Tongkui Yu[1,2] and Shu-Heng Chen[1] *

1. AI-ECON Research Center, Department of Economics, National Chengchi University, Taipei 11605
2. Department of Systems Science, School of Management, Beijing Normal University, Beijing 100875

Abstract. We propose a simple agent-based model of the political election prediction market which reflects the intrinsic feature of the prediction market as an information aggregation mechanism. Each agent has a vote, and all agents' votes determine the election result. Some of the agents participate in the prediction market. Agents form their beliefs by observing their neighbors' voting disposition, and trade with these beliefs by following some forms of the zero-intelligence strategy. In this model, the mean price of the market is used as a forecast of the election result. We study the effect of the radius of agents' neighborhood and the geographical distribution of information on the prediction accuracy. In addition, we also identify one of the mechanisms which can replicate the favorite-longshot bias, a stylized fact in the prediction market. This model can then provide a framework for further analysis on the prediction market when market participants have more sophisticated trading behavior.

Keywords: Prediction market, Agent-based simulation, Information aggregation mechanism, Prediction accuracy, Zero-intelligence agents, Favorite-longshot bias

1 Introduction

Prediction Markets, sometimes referred to as "information markets," "idea futures" or "event futures", are markets where participants trade contracts whose payoffs are tied to a future event, thereby yielding prices that can be interpreted as market-aggregated forecasts [1]. To predict whether a particular event (say, some candidate winning the election) will happen, a common approach is to create a security that will pay out some predetermined amount (say, 1 dollar) if the event happens, and let agents trade this security until a stable price emerges; the price can then be interpreted as the consensus probability that the event will happen. There is mounting evidence that such markets can help to produce forecasts of event outcomes with a lower prediction error than conventional forecasting methods [2].

* Corresponding author: chen.shuheng@gmail.com

To explain the efficiency of the prediction market in relation to aggregate information, many researchers use the efficient market hypothesis which attributes the market efficiency to a pool of knowledgeable traders who are capable of setting prices and acting without bias [3]. While Manski proposed a model [4] to find that price is a particular quantile of the distribution of traders' beliefs, price does not reveal the mean belief that traders hold, but does yield a bound on the mean belief. It can explain both the market efficiency and another stylized fact in prediction markets - *favorite-longshot bias* - which means that likely events (favorite) are underpriced or underestimated and unlikely events (longshot) are overpriced or overestimated [5]. Wolfers and Zitzewitz provided sufficient conditions under which prediction market prices coincide with average beliefs among traders in a model with log-utility agents [6]. Snowberg and Wolfers found evidence that misperceptions of probability drive the favorite-long shot bias, as suggested by prospect theory [7].

In this paper, we follow the recent research trend in agent-based prediction markets, and construct a spatial agent-based political futures markets based on the two-dimensional cellular automata. We begin this study with the device of zero-intelligence agents which was introduced into agent-based economic modeling by Gode and Sunder [8] and later on was applied to prediction markets by Othman [9]. However, instead of studying the general-purpose prediction markets, we focus on the political futures market, which, needless to say, is one of the most active application areas of prediction markets. This focus motivates us a spatial extension of the Othman's model. This extension also enables us to address a number of issues which cannot not be easily approached by either the neoclassical models of prediction markets [4, 7] or by agent-based prediction markets without spatial configurations. Specifically, the question is how exactly the information dissemination affect the information aggregation given that agents can only form their beliefs based on their information from their surroundings. Second, to take into account the geographical or social segregation phenomena, as analyzed by Schelling[10], we also study how clusters and their size may affect the operation efficiency of the political future markets. Using this Schelling-lik model, we can study the effect of cluster size to the prediction accuracy of the political future markets.

The rest of the paper is organized as follows. Section 2 introduces our spatial agent-based model of prediction markets, built on the very simple behavioral assumption of agents, namely, the zero-intelligent agents. Section 3 presents the agent-based simulation results. Section 3.1 shows that, by very simple behavioral assumptions of the traders, our agent-based model can replicate the well-known favorite-longshot bias. Section 3.2 further shows the prediction accuracy in monotonically (linearly) increasing in terms of the neighborhood size. Section 3.3 studies the effect of the geographical distribution of information on prediction accuracy. We find that, given the neighborhood size, there is a non-linear relation between block size and prediction error. Section 3.4 searches for the possible origins of the favorite-longshot bias, and indicates how the bid and ask

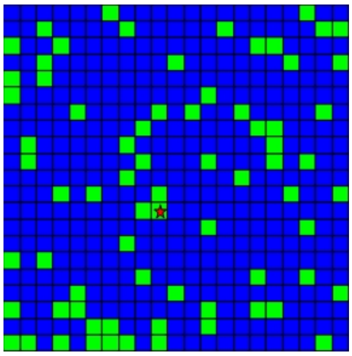


Fig. 1. Illustration of voter's political preference in a two-dimensional grid.

behavior can attribute to the emergence of this bias. Concluding remarks are given in Section 4.

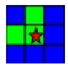
2 Basic Model

In a society represented by a two-dimensional torus grid, each element is a person, and each person has his own disposition on the vote for some candidate, blue for supporting (1) and green for not supporting (0) as in Figure 1. Figure 1 is simply for the illustrative purpose. In our simulation, we consider a more extensive model with a grid size of 200×200. Furthermore, in the simulations, the agents' voting disposition is randomly initialized with a given overall support ratio. This parameter is denoted as $SupportRatio$.

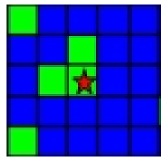
Some agents randomly sampled from the entire population will participate in a prediction market which provides a winner-take-all contract that pays a dollar if the candidate wins, and pays nothing if the candidate loses.[1] Each agent has a belief (subjective probability) $b_i \in [0, 1]$ that a candidate will win. This subjective belief is formed based on the sample statistics (sample mean) of the voting disposition of agent's neighbors. His/her neighbors consist of the center node (the agent himself) and the nearest Moore neighbors. Hence, the neighborhood has 9 agents for neighborhood radius $r = 1$, and 25 agents for neighborhood radius $r = 2$, as in Figure 2. For the agent marked with the red star, his own vote disposition is to support the candidate, and there are two other agents in his neighborhood with radius $r = 1$ who also support the candidate (Figure 2, left panel), so his belief is $b_i = 3/9 = 1/3$. If the radius $r = 2$ (Figure 2, right panel), his belief is $b_i = 5/25 = 1/5$.

This belief (subjective probability) will be taken as the reference price in the following sense that the agent would like to sell the future with any price

[1] The winner-take-all market and the share market are the two commonly used designs for the prediction markets. For the latter, the market participant is paid, according to the election results, by the final voting share of the candidate. Our agent-based model introduced here is equally applicable to the share market.



(a) r=1

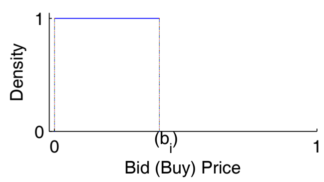


(b) r=2

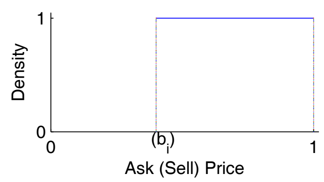
Fig. 2. Moore neighborhood of the agent in a two-dimensional grid

higher than this one, and would like buy with any price lower than this one. More specifically, by using the device of the *zero-intelligence agent*, the agent with belief b_i places a bid order for one share of the event at a price *uniformly* on $[0, b_i]$, or an ask order for one share at a price *uniformly* on $[b_i, 1]$, as in figure 3. The role of the agent, a buyer (to bid) or a seller (to ask), is randomly determined with equal probability. In vein of the zero-intelligence design, the learning or strategic behavior of agents is not taken into account. Agents do not "observe" (care) current market prices, and do not react to the result of their previous actions; they keep no record of previous unfinished orders.

This is a device of the *zero-intelligence agent* initiated by Gode and Sunder [8], which is now widely used in the agent-based models. The zero-intelligence agent is a randomly-behaving agent or, more precisely speaking, an entropy-maximizing agent. Since normal traders would not propose or accept a deal which would obviously lead to economic loss or not lead to welfare improvement, under no further information on what else they will do, the design of the zero-intelligence agent is *minimally prejudiced* in the sense of entropy maximization. The *uniform distribution* is employed here to realize the maximum entropy. Othman also carried out this design in his pioneering study on the agent-based prediction market [9].



(a)



(b)

Fig. 3. Order price formation

Following [9], the transactions are closed by the mechanism of continuous double exchange. Agents place buy or sell orders continuously. Once the highest-priced bid exceeds the lowest-priced ask, a trade occurs at the price of the order which was placed first. The paper, however, differs from [9] by explicitly embedding the agent-based prediction market within the network structure, a checkerboard as demonstrated in Figure 1. We consider this as a first attempt to hybridize the spatial agent-based political election models and the agent-based prediction markets for political elections.

The procedure of the model is described by the pseudocode in Algorithm 1.

Algorithm 1 Simulation Procedure

Require: Total number of agents N; Rounds of the market M; The overall support ratio *SupportRatio*;
Ensure: Transaction price history $TransactionPrice$

```
// INITIALIZATION
Generate a random number rndNum uniformly from 0 to 1
for each agent in the population do
  if rndNum < SupportRatio then
    Set his voting disposition as supporting (1)
  else
    Set his voting disposition as not supporting (0)
  end if
end for
// RUNNING THE MARKET
for round ∈ [1, M] do
  Choose an agent randomly from the whole population
  Get the voting dispositions of his neighbors
  Set the agent's belief b as the number of supporters over neighborhood
  size
  Generate a random number rndNum uniformly from 0 to 1
  if rndNum < 0.5 then
    Set the order side to 1 (buy)
    Set the OrderPrice as a random number uniformly drawn from 0 to b
    Get the MinSellPrice in the SellOrderList
    if OrderPrice > MinSellPrice then
      Insert a transaction in the TransactionList with MinSellPrice
    else
      Insert a buy order with OrderPrice in the BuyOrderList
    end if
  else
    Set the order side to 0 (sell)
    Set the OrderPrice as a random number uniformly drawn from b to 1
    Get the MaxBuyPrice in the BuyOrderList
    if OrderPrice < MaxBuyPrice then
      Insert a transaction in the TransactionList with MaxBuyPrice
    else
```

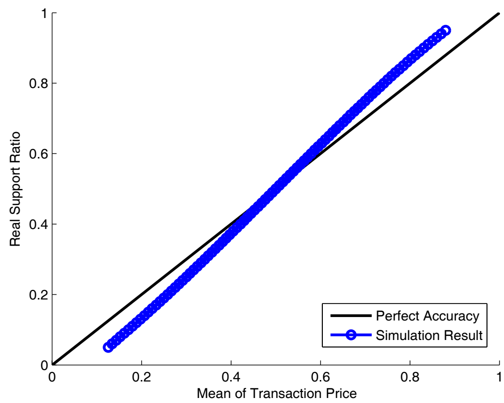


Fig. 4. Typical simulation result and favorite-longshot bias.

Insert a sell order with $OrderPrice$ in the $SellOrderList$
end if
end if
end for
end

3 Simulation and Analysis

We perform a great deal of simulations and try to find the regularity in the simulated data.

3.1 Prediction power and favorite-longshot bias

In the agent-based model of the political election prediction market, the mean of the transaction price is normally taken as a good predictor of the real support ratio in the overall population. Figure 4 provides the relationship between the real support ratio (vertical axis) and the mean of the transaction price (horizontal axis) in typical simulations with the neighborhood radius $r = 1$. This simulation is conducted in a 200 by 200 checkerboard with 40,000 agents ($N = 40,000$), one in each cell (a checkerboard with full size). For each given support ratio, the continuous double auction is run once for 40,000 rounds. We then try support ratios from 0.1 to 0.9 with an increment of 0.01; in other words, a total of 81 support ratios are tried. We then plot the mean of the transaction price over the 40,000 rounds for each support ratio in Figure 4. As we can see from that figure, the mean transaction price can trace the true support ratio to some degree.

At the same time, the simulation replicates the favorite-longshot bias, a stylized fact in the prediction market. We can find that unlikely events (bottom-left

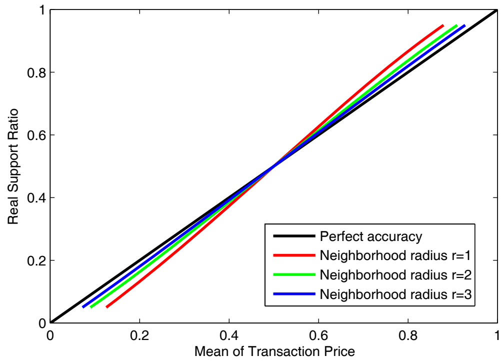


Fig. 5. Neighborhood radius and prediction accuracy.

in Figure 4) are overpriced, and likely events (top-right in Figure 4) are underpriced.

3.2 Neighborhood scope and prediction accuracy

We investigate the effect of the neighborhood scope on the prediction accuracy. Figure 5 provides the relationship between the real support ratio (vertical axis) and the mean of the transaction price (horizontal axis) in typical simulations with different neighborhood radii $r = 1, 2$ and 3. We can find that the prediction accuracy of the prediction market increases as the neighborhood scope increases. It is easy to understand that the more information that each participant has, the more accurate the prediction that the market can provide. This result is similar to the work of Othman [9]. However, we obtain the result with a totally different basic assumption. We use rather simple assumption that the agent forms his belief by observing his neighbors' voting deposition, while Othman's work requires specific distribution of belief. With this distinction, we attribute the extent of the bias to the amount of individual information that each individual agent has, while Othman's work attributes it to the arbitrarily specified belief distribution.

To measure the prediction accuracy of a prediction market, we define a variable referred to as the mean squared error

$$\rho = \frac{\sum_1^S (MeanPrice_i - SupportRatio_i)^2}{S}, \tag{1}$$

where $i = 1 : S$ is the number of simulations. A larger mean squared error implies less accurate prediction, while a smaller one implies more accurate prediction, and $\rho = 0$ implies perfect accuracy. With the same neighborhood radius, we simulate $S(S = 200)$ times with $SupportRatio$ linearly spaced between 0 and 1 and calculate the prediction accuracy. Figure 6 presents the effect of the neighborhood radius r (horizontal) on the prediction accuracy (vertical axis). The

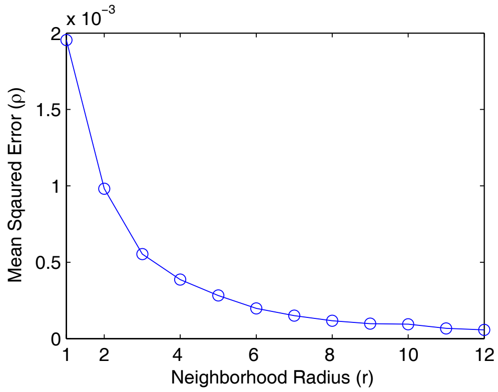


Fig. 6. Effect of neighborhood radius on the prediction accuracy.

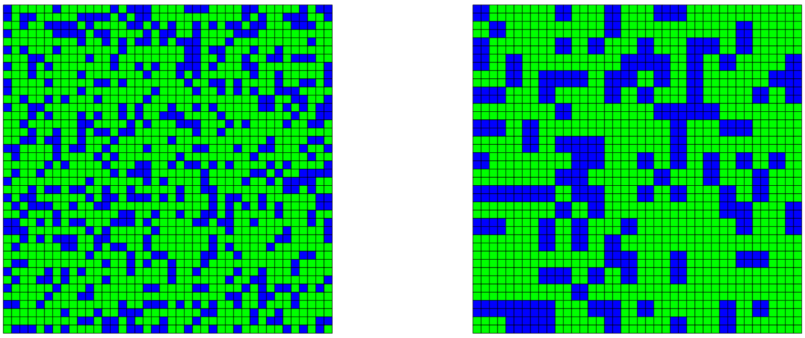
larger the neighborhood radius r, the more information agents have, and the more accurate prediction the market can provide.

3.3 Information distribution and prediction accuracy

In a real political election, the voting deposition may be clustered. Some places may be dominated by the supporter of one candidate, and most of the population may support the other candidate elsewhere. So the information is not well scattered. We wish to study the effect of information distribution on the prediction accuracy. To this end, we manipulate the block size of voting disposition to represent different information distributions.

When initializing the voting disposition of the agents, we use different granularities or block size s. If the block size $s = 1$, we initialize the agents' voting disposition one by one; if the block size s=2, we initialize the agents' voting disposition two by two, i.e., all four agents in the 2×2 sub-grid have the same voting disposition randomly generated according to the specified support ratio. Figure 7 illustrates a comparison of information distributions with block sizes $s = 1$ and 2 under the same support ratio 0.3.

Figure 8 depicts the effect of voting disposition block size on prediction accuracy with neighborhood radii $r = 2, 3, 5$ and 7. We can find the nonlinear relationship between the block size and mean squared error. The mean squared error is the largest when the voting disposition block size is close to the neighborhood radius, and the market prediction power is the least accurate. The farther away the voting disposition block size is from the neighborhood radius, the smaller is the mean squared error, the more accurate is the market prediction. This result is confirmed by figure 9 which provides the combination effect of block size and neighborhood radius to the mean squared error. Further research is needed to understand this phenomenon.



(a) $s = 1$ (b) $s = 2$

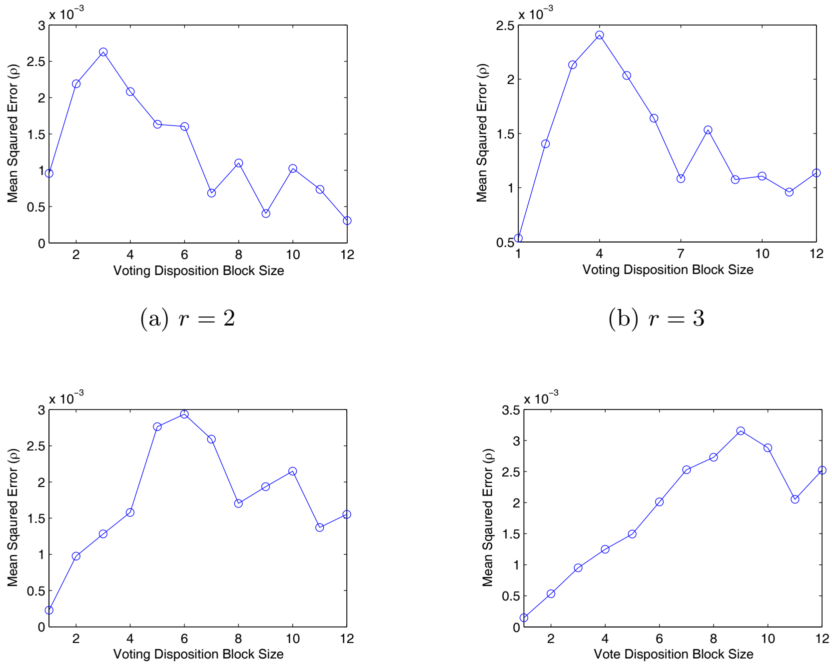
Fig. 7. Comparison of information distributions with different block sizes of voting disposition

3.4 The origin of the favorite-longshot bias

So far, we have shown that we can replicate the favorite-longshot bias. The advantage of using agent-based models is that we can go further to ask whether we can trace the possible source of this bias. To find the mechanism that produces the favorite-longshot bias, we have performed some experiments. One possible origin of the favorite-longshot bias is the *price formation mechanism*; the baseline version closes transactions at the price of the order (bid or ask) which was first placed. An alternative price formation mechanism is to take average of the bid and ask, instead of only one of the two (the earlier posted one). Nonetheless, our simulation shows that the favorite-longshot bias still exists with this modification.

The other possible origin of the favorite-longshot bias is the *order-formation mechanism*. The baseline version is that an agent can submit a bid between zero and his belief b_i or an ask at a price between b_i and 1 (as in Figure 3). When the agents' belief is not equal to 0.5, there is an inherent asymmetry between the range of bid and the range of ask. This may lead to the favorite-longshot bias.

We test the hypothesis by proposing a symmetric order price formation mechanism where an agent submit a bid at a price drawn randomly from $[b_i - \delta, b_i]$, or submit an ask at a price drawn randomly from $[b_i, b_i + \delta]$ as in Figure 10(a) . We find that the favorite-longshot bias disappears in this specification as in 10(b). Moreover, we perform a simulation using the designed asymmetric order price formation mechanism, where the prices of buy orders are uniformly on $[b_i - \delta, b_i]$ and the prices of sell orders are uniformly on $[b_i, b_i + 2\delta]$ as in Figure 10(c), and we find that the prices are all overvalued as in figure 10(d). Furthermore, if the prices of buy orders are uniformly on $[b_i - 2\delta, b_i]$ and the prices of sell orders are uniformly on $[b_i, b_i + \delta]$ as in Figure 10(e), the prices are all undervalued as



(a) $r = 2$ (b) $r = 3$

(c) $r = 5$ (d) $r = 7$

Fig. 8. The effect of voting disposition block size on prediction accuracy

in Figure 10(f). So we can come to the conclusion that the asymmetry of the order-price mechanism is the one of possible origins of the favorite-longshot bias.

4 Conclusion

Prediction markets and experimental markets are, so far, the two real tests for the Hayek hypothesis on the market mechanism as functions of information aggregation and information externality. Given the competitiveness and the popularity of the ideas of prediction markets, it would be imperative to see how this idea actually works, and agent-based modeling can provide possible rich settings to examine their functions. In this specific paper, we show how the well-known favorite-longshot bias can be replicated through one version of our agent-based model; but we also show how different settings may cause this bias to disappear. Neighborhood size and cluster size all have their effects on the accuracy of the prediction markets, but the cross interaction of the two needs to be further examined.

One extension of this paper is to replace the zero-intelligence agents with agents behaving more realistically. This is the part which behavior finance may

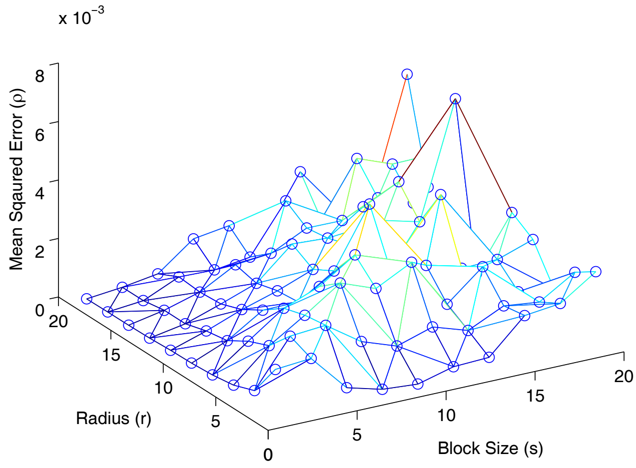


Fig. 9. Block Size, neighborhood radius and prediction accuracy.

shed light on. As what has been analyzed in this paper, the favorite-longshot bias can be caused by the asymmetric trading behavior of the agent, and the bias disappears when symmetric trading behavior is imposed. However, regardless of being symmetric or asymmetric, the zero-intelligence trading behavior characterized by the uniform-distributional bids or asks is not realistic, at least for economists.[2] Therefore, one could argue that the favorite-longshot bias can equally likely be caused by more deliberate and sophisticate trading behavior. If so, what would be the minimal description of the "smart" behavior leading to this bias is a question for the future of the research. Finally, whether the simulation results can have important implications for the design of the prediction market is also an issue for further study [11].

Acknowledgements

The financial support from the Chung-Hwa Development Foundation and the NSC grant 98-2410-H-004-045-MY3 are gratefully acknowledged.

References

1. Wolfers J., Zitzewitz E.: Prediction Markets. J. Econ. Perspect. 18(2), 107-126 (2004)

[2] For example, as one of the referees has correctly pointed out that our zero-intelligence design means that an agent that believes the share is worth 0.5 is equally likely not to sell it for 0.6 or 0.9.

2. Arrow K. J., Forsythe R., Gorham M., et al.: The Promise of Prediction Markets. Science. 320, 877-878 (2008)
3. Forsythe R., Nelson F., Neumann G., Wright J.: Anatomy of an Experimental Political Stock Market. Am. Econ. Rev. 82(5), 1142-1161 (1992)
4. Manski C. F.: Interpreting the Predictions of Prediction Markets. Econ. Lett. 91(3), 425-429 (2006)
5. Wolfers J., Zitzewitz E.: Five Open Questions About Prediction Markets. In: R. W. Hahn (eds.) Information Markets: A New Way of Making Decisions. AEI Press (2008)
6. Wolfers J., Zitzewitz E.: Interpreting Prediction Market Prices as Probabilities. NBER Working Paper No. 10359 (2005)
7. Snowberg E., Wolfers J.: Explaining the Favorite-Longshot Bias: Is it Risk-Love or Misperceptions? J. Polit. Econ. 118(4), 723-746 (2010)
8. Gode D., Sunder S.: Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality. J. of Polit. Econ. 101(1), 119-137 (1993)
9. Othman A.: Zero-intelligence Agents in Prediction Markets. In: Proceedings of AAMAS 2008, pp. 879-886 (2008)
10. Schelling T (1971) Dynamic models of segregation. Journal of Mathematical Sociology 1:143-186.
11. Chen S-H., Tung C-Y., Tai C-C., Chie B-T., Chou T-C., Wang S. (2011) Prediction Markets: A Study on the Taiwan Experience. In: Williams L. (ed.), Prediction Markets: Theory and Applications, London: Routledge. Chapter 11, pp. 137-156.

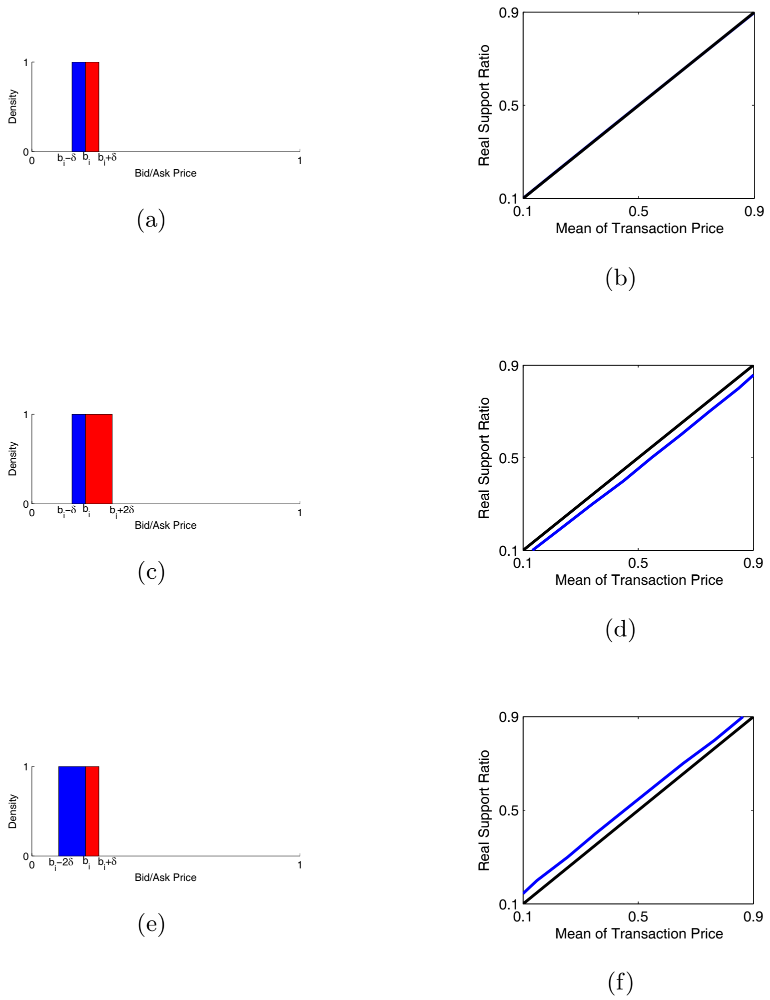


Fig. 10. Order price formation mechanism and favorite-longshot bias