The Use of Knowledge in Prediction Markets: How Much of Them Need He Know?*

BIN-TZONG CHIE[1] AND SHU-HENG CHEN[2]
[1]*Department of Industrial Economics*
Tamkang University
Tamshui, New Taipei City, 251 Taiwan
E-mail: chie@mail.tku.edu.tw
[2]*AI-ECON Research Center*
Department of Economics
National Chengchi University
Taipei, 116 Taiwan
E-mail: chen.shuheng@gmail.com

In this article, we extend an early agent-based spatial model of the prediction market by taking into account the heterogeneities of agents in their tolerance capacity (tolerance to neighbors with different political identities) and in their exploration capacity (exploration of the political identities of other agents). We then study the effects of these heterogeneities on the behavior of the prediction market, including prediction accuracy, determinants of earnings, and income distribution. First, in terms of prediction accuracy, we find that, compared to the homogeneous case, bringing heterogeneity into the model can generally improve the prediction accuracy, although its statistical significance is limited. In particular, the well-known empirical regularity known as the favorite-longshot bias remains almost unchanged with this extension. Second, through the heterogeneous-agent design, we find that both capacities (personality traits) of agents have a significant positive effect on earnings, and the effect of the exploration capacity is even more dramatic. Third, through their effects on earnings, both capacities also contribute to income inequality, but only to a mild degree with a Gini coefficient of 0.20.

Keywords: prediction markets, Schelling segregation model, zero-intelligence agents, Hayek hypothesis, tolerance capacity, exploration capacity, favorite-longshot bias

1. MOTIVATION AND INTRODUCTION

How much knowledge does he need to do so successfully? Which of the events which happen beyond the horizon of his immediate knowledge are of relevance to his immediate decision, and how much of them need he know? ([1], p. 525)

In his celebrated article "The Use of Knowledge in Society" [1], Hayek argued that the price system provides the solution to the use of knowledge required to solve real problems. Having said that, Hayek did not consider that there was any better alternative which could help the decision maker(s) to acquire the knowledge required to solve the problem. While this paper has been read out loud on many occasions in classrooms, meeting rooms, journal articles, and now even over the Internet, and has been embraced

Received December 10, 2013; revised August 20, 2014; accepted October 24, 2014.
Communicated by Chao-Lin Liu, Yung-Jen Hsu, Shou-De Lin, Kuo-Wei Hsu and Ming-Feng Tsai.
* The paper was originally presented at the *2013 Conference on Technologies and Applications of Artificial Intelligence* at National Chengchi University, Taipei, December 6-8, 2013, and was substantially revised and expanded by fully taking into account of three anonymous referee reports. The authors are grateful to the three referees for their painstaking reviews of the paper. All remaining errors are authors' sole responsibility. The NSC grant NSC 101-2410-H-004-010-MY2 is also gratefully acknowledged.

by uncountable celebrities as part of their wisdom, it remains unclear to us as to what is the necessary knowledge required for an individuals to know or, in Hayek's own words, "*how much of them need he know?*"

As far as we can see, there is a *co-dependence issue* in Hayek's argument. From beginning to end, Hayek very frequently seemed to emphasize that what individuals need to know is very limited given the support of a price system; nonetheless, he seemed to say little to remind us that the price system is in the meantime constituted by individuals with different degrees of acquired knowledge. If the individual becomes more ignorant due to a good awareness of the price system, for example, only making his decision by looking at the Dow Jones index and not reading the Wall Street Journal, and everyone else behaves in a similar manner, then the quality of the emergent price system may not warrant that good decisions can be made by watching "merely the movement of a few pointers" (Ibid, p. 527)?[1]

Maybe in Hayek's grand analysis, this co-dependence issue is rather minor, particularly considering that at the time when the paper was written, he was defending the market economy against the centrally planned economy.[2] However, the "minor" issue may become serious if we misread the paper by taking the price system for granted and ignoring individuals' roles in the aggregation process.

Half a century later, economists started to apply the Hayek's assertion to institutionalize a price system, known as the *prediction market*, by which the knowledge distributed widely among individuals can be aggregated and used to predict future events. It is in this practical application that Chie and Chen [2] examined the *role of individuals* in Hayek's assertion, also referred to as *Hayek's hypothesis* by Vernon Smith [3]. They developed a spatial agent-based model of the prediction market. They first applied Schelling's segregation model [4] with a given tolerance level, s, also known as the migration parameter, to mimic the clustering phenomenon driven by different political identities.[3] The issue of limited knowledge or imperfect knowledge as addressed by Hayek is then characterized by agents' exploration capacity (r), which, physically speaking, just determines a neighborhood size of agents. Essentially, this parameter determines how much agents endeavor to know, and hence is a measure of the efforts devoted to knowledge acquisition. This setup allows us to examine the role of individuals in knowledge acquisition in the Hayek hypothesis.

In [2], the two capacities, the tolerance capacity (s) and the exploration capacity (r), are assumed to be identical for all agents in the society. What the authors did was to distinguish one society from the other society by manipulating these two parameters. Hence, the study seeks to examine the performance of the prediction market over different societies characterized by these two capacities. Surprisingly, it was found that while beefing

---

[1] Vriend [5] studied the aggregation problem in Hayek's hypothesis under different behavioral settings. Using an agent-based model, he was able to distinguish between the cases where the information was aggregated well and the cases where it was not. When the social medium has gained its dominant role in human communication and decision-making processes, the worry of *the stupidity of the herd* is also increasing, no less than the appraisal for the *wisdom of crowds* [6, 7].

[2] This is known as the socialist calculation debate [8].

[3] What considered in [2] is a spatial cluster, which is used to approximate the political map normally associated with election outcomes. However, the cluster phenomenon is not limited to the spatial configuration, but is generally applied to social networks as well. We have not done this part yet, which shall be an issue left for further study.

up these two capacities can make the society have better-informed individuals, it does not enhance the performance of the prediction markets. On the contrary, it makes it worse. This seemingly counter-intuitive result is due to the *homogeneity effect*. Basically, enlarging these two capacities actually enhances the flow of the information over various corners of the society, which causes agents to become rather homogeneous in their beliefs and reservation prices and hence results in bids and asks being difficult to meet. This result is what is familiarly known as the *no-trade theorem* [9].[4]

In this paper, we address the significance of *heterogeneity*. We assume that agent *i*'s capacities are stochastically determined as follows:

$$s_i = s + \varepsilon_s, \quad \varepsilon_s \sim U[-a, a], \tag{1}$$

and

$$r_i = r + \varepsilon_r, \quad \varepsilon_r \sim U[-b, b]. \tag{2}$$

Hence, we denote this new heterogeneous scenario as $\{s, r; \varepsilon_s, \varepsilon_r\}$, and the original homogeneous scenario as $\{s, r; 0, 0\}$ or simply $\{s, r\}$. In this way, the two scenarios are identical in mean (certainty equivalence), but differ in their distribution. Therefore, by simulating both scenarios and comparing their results, we can evaluate the heterogeneity effect or the *distribution effect.* Specifically, we ask the question, *under the certainty-equivalence condition, whether the prediction market can perform better in the heterogeneous scenario than in the homogenous scenario, and, if so, how much better*? This is the key research question to be addressed in this paper.

Obviously, taking the distribution effect into account does not go against the dispersion of knowledge among different individuals; on the contrary, it is probably even closer to what Hayek [1] had assumed. As we shall elaborate later, doing this allows us to introduce *personality traits* into the model, and personality traits by definition are heterogeneous among agents. Because of this addition, this paper can also be read as a pioneering step in addressing the effect of personality on information aggregation or the use of knowledge in the prediction market.

The rest of this paper is organized as follows. Section 2 will introduce our spatial agent-based model of the prediction market. Section 3 gives the simulation design. Section 4 discusses the simulation results. Section 5 gives the concluding remarks.

2. THE MODEL

The spatial agent-based model of the prediction market proposed in this paper is extended from Yu and Chen [10]. Early agent-based prediction markets did not take into account an embedded social network in which market participants obtain their personal information. Yu and Chen [10] made a first step toward it by bringing the idea of the political map into the model. However, the kind of clusters which they introduced was square blocks. Hence, the geographical features so formed are rather distant from the general political map observed in reality. Chie and Chen [2] first applied Thomas Schel-

[4] Chie and Chen [6], however, did find that, after controlling the homogeneity effect, the two capacities have a positive effect on market performance.

ling's segregation model to generate the kind of irregular clusters which may fit the reality better. This spatial model has three mainstays, namely, a trading institution (section 2.1), a physical (spatial) network (section 2.2), and a social network (section 2.3).

2.1 Trading Institution

Network-Based Formation of Expectations and Reservation Prices We shall make both the spatial and social networks explicit in our model (Sections 2.2 and 2.3). Through the given spatial and social networks, agents disseminate and acquire the information and form their expectations of the future of election outcomes, upon which their decisions on bids and asks are based. We assume that, to form an expectation of the election outcome, all agents use the sample average as the estimate, and the sample available to each agent is identical to the set of all his connecting agents (to be defined in Section 2.3). In other words, by using the sample proportion of the connecting agents supporting each political candidate, the agents form their expectations about the share of the vote of each candidate. This estimated share becomes the *reservation price* hold by the agents. To make this point precise, let $\hat{p}_{i,j}$ be the subjective estimation of agent i regarding the share of the votes attributed to candidate j, and $b_{i,j}$ be the reservation price that agent i holds for the futures related to the vote share of candidate j. Then

$$b_{i,j} = \hat{p}_{i,j} = \frac{\#\{k : k \in N_i \cap V_j\}}{\# N_i}, \quad i = 1, 2, ..., N, \quad j = 1, ..., m, \tag{3}$$

where N_i is the set of agent i's connecting agents (to be defined in Section 2.3), and V_j is the set of voters who support candidate j. For example, by (3), if the estimated share of the votes of Candidate A is 60%, then the reservation price of the future contract for Candidate A is 60 cents. With this reservation price, the agent would not accept any bids which are lower than 60 or any asks which are higher than 60.

Bidding and Asking Strategy In fact, following most agent-based prediction markets [10, 11], we assume that all agents are *zero-intelligent agents* (the entropy-maximizing agent) in the sense that the agent will bid or ask randomly with the constraint of making no expected loss [12, 13]. Therefore, his bid $p_{b,i,j}$ will be uniformly sampled from the interval between the floor, which is zero cents, and the reservation price $b_{i,j}$, and his ask $p_{a,i,j}$ will be uniformly sampled from the interval between his reservation price and the ceiling, which is one dollar, as shown in Eq. (4).

$$p_{b,i,j} \sim U[0, b_{i,j}], \quad p_{a,i,j} \sim U[b_{i,j}, 1], \quad i = 1, 2, ..., N, \quad j = 1, ..., m. \tag{4}$$

Double Auction The trading mechanism adopted to run the market is continuous double-auction, the one frequently used in the experimental economics to test the Hayek hypothesis [3]. The mechanism has also been used in a number of other agent-based prediction markets [10, 11]. Fig. 1 gives a summary of the flow of the mechanism. The market starts from a random draw of the agents. Each agent shall be drawn exactly once; in other words, the draw proceeds in a sampling-without-replacement manner. When agent i is drawn, he will be randomly placed into one of the m markets and will be

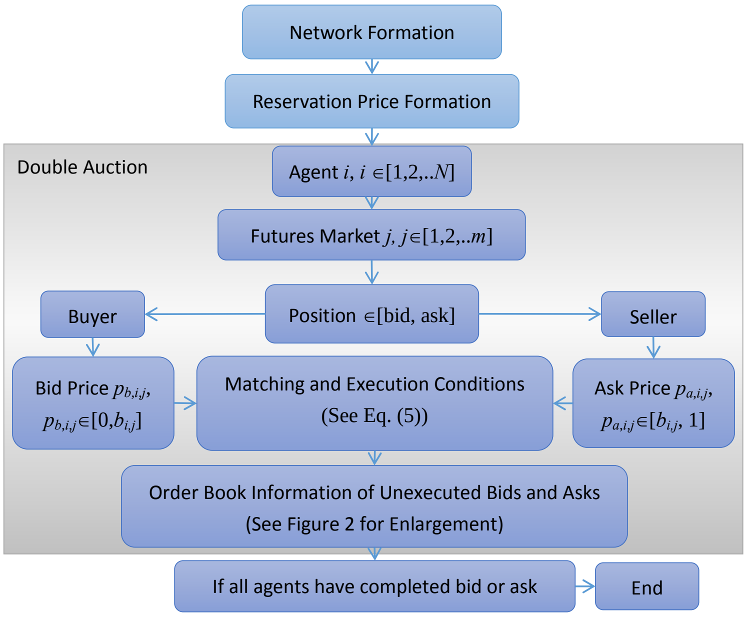


Fig. 1. The flowchart of the order book-driven prediction market.

equally likely assigned either as a buyer position or a seller position. He will then submit a bid if he is a buyer and submit an ask if he is a seller. His bid or ask will be placed in the *order book*. A match happens if either his bid ($p_{b,i,j}$) is no less than the remaining lowest ask (bestp_a) in the order book or his ask ($p_{a,i,j}$) is no greater than the remaining highest bid (bestp_b). The transaction price will then be determined as bestp_a if the former applies or as bestp_b if the latter applies.

$$p = \begin{cases} \text{best}p_a, & \text{if } p_{b,i,j} \geq \text{best}p_a, \\ \text{best}p_b, & \text{if } p_{a,i,j} \leq \text{best}p_b. \end{cases} \tag{5}$$

The flow of the double auction market is illustrated in Fig. 2. The figure is divided into three panels. The left panel shows the event time (the first column); the middle panel gives the information of traders, including their reservation prices (the third column), which is determined by their connecting agents in a way as indicated in Eq. (3); the right panel shows that the constantly updated order book information of unexecuted bids and asks (the sixth and seventh columns) as well as all transaction prices (the last column).

In this specific example, a number of agents are sequentially randomly sampled without replacement to enter the market; as shown in the second column, they are 9003, 3882, 9152, *etc.* They are also randomly assigned a position, either buyer (to bid) or seller (to ask), as shown in the fourth or the fifth column. Given their reservation prices and positions, they bid or ask according to Eq. (4), and the realizations are shown in the fourth or fifth column. Given the bids and the asks, the market constantly checks the matching condition (Eq. (5)), executes those matched bids and asks, releases the transaction price (the last column), and updates those unexecuted bids and asks (the sixth and

Event	Information of Agents				Futures Market of Blue Candidate		
	Agent	**reservation price**	**bid/buyer**	**ask/seller**	**Bid**	**Ask**	**Price**
1	9003	0		**47.03**		**47.03**	-
2	3882	57.14		**85.19**		47.03 **85.19**	-
3	9152	0		**100**		47.03 85.19 **100**	-
4	1569	64.71	**48.22**		**48.22**	47.03 85.19 100	-
5	match 48.22 >= **47.03**		buyer 1569	seller 9003		85.19 100	**47.03**
6	10949	56.21	**52.07**		**52.07**	85.19 100	47.03
7	8307	40.74	**29.83**		52.07 **29.83**	85.19 100	47.03
8	8686	51.16		**65.87**	52.07 29.83	**65.87** 85.19 100	47.03
9	576	87.47	**27.12**		52.07 29.83 **27.12**	65.87 85.19 100	47.03
10	3855	0		**25.38**	52.07 29.83 27.12	**25.38** 65.87 85.19 100	47.03
11	match 25.38 <= **52.07**		buyer 10949	seller 3855	29.83 27.12	65.87 85.19 100	**52.07**
12	… …	… …	… …	… …	… …	… …	… …

Note: The left panel shows the event time. The middle panel shows the event characterized by a random draw of an agent and his/her randomly determined position (bid or ask) or a match and an execution. The right panel presents the constantly updated order book, including all unexecuted bids and asks and all transaction prices thus far.

Fig. 2. An illustration of the flow of the double auction market with an order book.

seventh columns). As shown in this example, up to event time 11, the market has already received 4 bids and 5 asks and is able to only match two of them, one at event time 5 and one at event time 11, with a price of 47.03 for the former and a price of 52.07 for the latter.

2.2 Spatial Network

The spatial networks considered in this paper are generated from the *Schelling segregation model* [4], in which the location of agents is determined by their *tolerance capacity* for agents with different political identities. In other words, we replace the ethnic heterogeneity of agents in the original Schelling model with their *political identity* ($j = 1, 2, ..., m$). Agents tend to reside in the place which is surrounded by the neighbors with the same political identity. Their tolerance of neighbors with different political identities is characterized by the parameter, *tolerance capacity* (s). The tolerance capacity is a ratio between 0 and 1. If the tolerance capacity is 0, then the agent cannot accept any other agents with different political identities as his/her neighbors. In other words, if he/she has one neighbor with a different political identity, then the agent will decide to migrate to another area which is susceptible to his/her tolerance capacity. On the other hand, if it is 1, then the agent will accept any number of neighbors with different political identities, even though he/she may be a minority in the community.

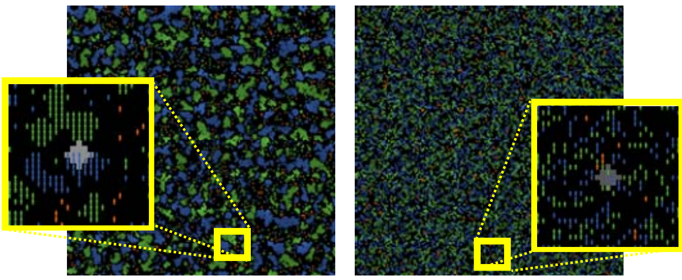
If the tolerance capacity is between 0 and 1, then the agent can accept the unlike agents up to a proportion of s. If the ratio of neighbors with different political identities is larger than this threshold s, they tend to move to a close place which their tolerance capacity can handle. This migration process will be iterated until it converges to a fixed configuration. We then use the resultant configuration to represent the geographical distribution of residents with different political identities.

Apart from the tolerance capacity, an additional parameter of Schelling's segregation model is the demographical structure characterized by the percentage of agents of various political identities. Denote them by v_j (j = 1, 2, ..., m).

$$v_j = \frac{\#(V_j)}{N},\ j = 1, 2, ..., m, \tag{6}$$

where N is the total number of agents.

Fig. 3 demonstrates a geographical distribution of political identities. In this specific example, there are a total of 13,454 agents, distributed on a checkerboard with 193 × 193 grids, *i.e.*, with a population density of 36.12%, and m = 3 (three candidates or three political parties): v_1 = 51.63%, v_2 = 45.63%, and v_3 = 2.77%. Agents having one of the three political identities are denoted by the blue (j = 1), green (j = 2), and the orange color (j = 3), respectively.[5] What is demonstrated in Fig. 3 is, therefore, two of the converged configurations of agents who followed the Schelling rule of migration. The one on the left is the one corresponding to a tolerance capacity of 0.25, and the one on the right is the one corresponding to a tolerance capacity of 0.75.



s = 0.25 s = 0.75

Note: Both panels are the converged configurations using v_1 = 45.63% (green), v_2 = 51.63% (blue), v_3 = 2.77% (orange), N = 13,454, and G (number of grids) = 193 × 193. The black grids denote the unoccupied cells, and the colored grids denote the occupied cells. The number of occupied cells and the number of unoccupied cells are determined in a way such that the resultant population density is close to 36% (see Table 1). The two panels differ in terms of the tolerance capacity: on the left, s = 0.25, and, on the right, s = 0.75.

Fig. 3. Geographical distribution of voters and their political identity.

[5] These parameter values are based on the 2012 Presidential Election in Taiwan. Based on the 2012 Presidential election outcome, the KMT candidate (colored in blue) won a share of 51.60% of the votes, the DPP candidate (colored in green) won a share of 45.63%, and the PFP candidate (colored in orange) won a share of 2.77%.

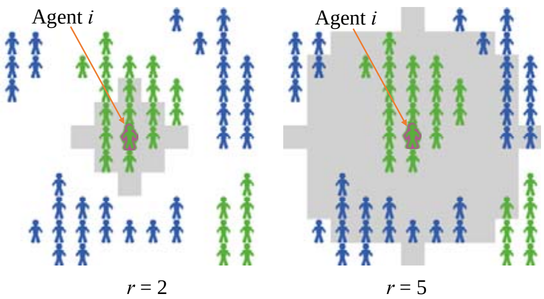
Clearly, as one can imagine, if the tolerance capacity is generally low, then by the Schelling migration process, the society tends to segregate into remarkably large clusters, *i.e.*, not just on the scale of a local community, but even on the scale of a village or a county in which a great majority of residents tend to have the same political identity. On the other hand, if the tolerance capacity is generally high, then it will be less easy to locate a sizable area with an overwhelmingly dominating political identity, of course, after the weights v_j (j = 1, 2, …, m) have been properly taken into account.

This description corresponds well to the two tolerance capacities of Fig. 3. The right panel, corresponding to s = 0.75, clearly has a highly integrated society in which the blue and the green are well mixed in every corner. However, the left panel, corresponding to s = 0.25 shows that emergence of many blue 'clusters' and green 'clusters'. The two panels are, therefore, in sharp contrast to each other. To see this contrast better, we zoom in to focus on specific areas of the map, and depict them in the lower left and lower right panels of the figure. The geographic distribution of the agents with different political identities then dictates how information is distributed spatially. Hence, the question concerning us is: Given such widely and unevenly distributed information, how well can the market actually aggregate them. This is the main concern of the Hayek hypothesis.

2.3 Social Network

While each agent is physically constrained to a limited geographical space, the agents' mobility is not necessary limited to this narrowly-defined neighborhood. Depending on their sociability, they may reach a set of agents which is larger than just their 'neighbors'. This is the place where the social network can come into play. Therefore, the third main stay of the model is to add a social network on the top of the Schelling segregation model.

Embedding a general social network within a physical network so as to observe the interplay of physical and social space on human behavior involves research which can be dated back to the late 1990s. The Schelling-Axelrod model of culture is one of the pioneering studies [14, 15]. Since the late 1990s, various social network topologies have been proposed and studied by computer scientists and physicists, which actually enable us to see the significance of topological characteristics to information flow and to various social and economic performances [16].



Note: The above figures shows the von Neumann neighborhood of agent i, as pointed by an arrow. The left panel is a neighborhood with a radius of 2, whereas the right panel is a neighborhood with a radius of 5.

Fig. 4. The von Neumann neighborhood with a radius of 2 (left) and 5 (right).

A social network can be introduced to the Schelling segregation model in many arbitrary ways. Nonetheless, for us, a more sensible way to do it is to leave agents to decide on the size of the social network by themselves based on their own characteristics. An intuitive way to implement this idea is to assume that different agents have different exploration capacities, which then determines the size of the network which they attempt to reach and maintain. Accordingly, we assume that, for each agent, his set of connecting agents is determined by a von Neumann neighborhood with a given radius (r). The radius, r, can be interpreted as the information exploration capacity of the agent. The larger the radius, the larger the sample is, and hence the less biased and the better the estimation. This is shown in Fig. 4.

Fig. 4 shows the social network of two agents who differ in terms of their exploration capacity. The agent with the smaller exploration capacity (r = 2), shown in the left panel, only has a near 'sight' of his neighbors (agents in the gray area), and, in this case, all the agents reached by him/her happen to be the same (green). The agent with the larger exploration capacity (r = 5), shown in the right panel, has a far 'sight' of his neighbors; hence, in addition to the green neighbors, he/she is also able to have a non-negligible sample of blue agents. If both agents use Eq. (3) to estimate the vote share for each candidate, then clearly the agent with the smaller radius may tend to have a larger bias than the one with the larger radius. This difference matters because, unlike Chie and Chen [2], in this article, agents are *heterogeneous* with respect to this capacity.

2.4 Programming with NetLogo

The above-mentioned spatial agent-based prediction market is programmed with NetLogo 5.1.0 and is available from the OpenABM website.[6] Fig. 5 shows a familiar NetLogo display of running this program.

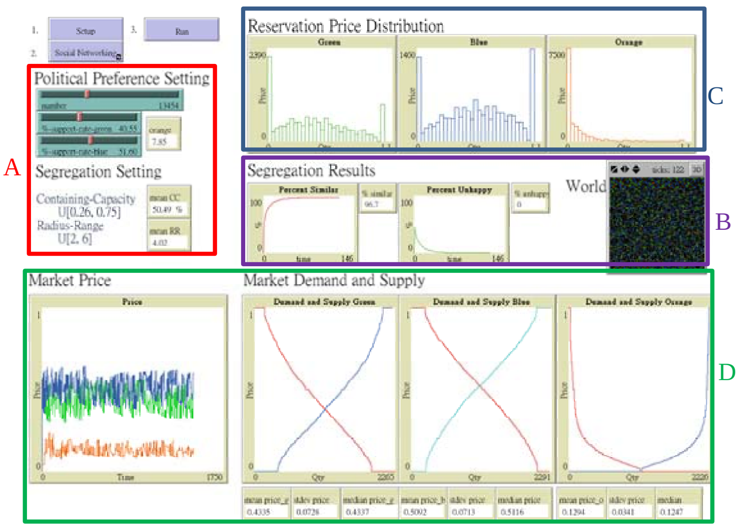


Fig. 5. Display of the NetLogo program.

[6] http://www.openabm.org/model/3764/

In Fig. 5, the upper left panel (panel A) gives the user-supplied control parameters: N = 13,454, v_1 = 40.55% (green), v_2 = 51.60% (blue), v_3 =7.85% (orange), s = 0.50 ($s_i \sim U$ [0.26, 075]) and r = 4 ($r_i \sim U$ [2, 6]). The diagram shown in the right middle panel (panel B) is the converged configuration using the Schelling rule with s = 0.5. With a mean radius of 4, we can have the price expectations (reservation prices) of all three futures for all agents, *i.e.*, $b_{i,j}$ (i = 1, ..., 13454, and j = 1, 2, 3). What is shown in the right upper panel (panel C) of the figure are the three histograms of the reservation prices corresponding to the green, blue and orange parties, respectively. The basic statistics, including the mean, the median and the standard deviation, are shown at the very bottom of the figure (panel D). There we can see that the mean and median for the green candidate is 0.4335 and 0.4337, which reflects a three-point (3%) upward bias away from the true value of 0.4055. In addition, for the blue candidate, these two statistics are 0.5092 and 0.5116, reflecting a downward bias away from the true value of 0.5160 of less than one point. Maybe the worst case is the market for the orange candidate. The two corresponding statistics are 0.1294 and 0.1247, almost two times larger than the true value of 0.0785. Our research question is then to what extent this specific network topology may affect the accuracy of the prediction market or the political futures market in our case.

From the histogram, we can further derive the aggregate willingness to buy (for those whose reservation price is less than the market price)

$$Q_j^D(p) = \#\{i : b_{i,j} < p\}, \tag{7}$$

and the aggregate willingness to sell (for those whose reservation price is greater than the market price)

$$Q_j^S(p) = \#\{i : b_{i,j} > p\}, \tag{8}$$

i.e., the demand curve (Q_j^D) and the supply curve (Q_j^S). The demand and supply curves of the three markets are shown in the lower middle and right panel (panel D). Then through the random draws of the agents and their reservation prices, the order book for each market is formed, and the corresponding transaction price is generated as the time series shown in the lower left panel (panel D) of the figure.

3. EXPERIMENTAL DESIGN

3.1 Number of Agents and Density

The setting of other parameters is basically the same as in Chie and Chen [2]. First, the number of agents and the population density are set in such a way that they, to some extent, can mimic the Taiwan demographic structure, after a proper scaling-down adjustment. The details are as follows.

According to the 2010 demographic census data in Taiwan, there were 7,414,327 households. By first scaling down the number of households by a factor of 1,000, we have a total of 7,414 households. We then assume that the space required for each household has a size of 5 grids. This assumption is based on the consideration that, for 93% of the households in Taiwan, the number of members has a maximum size of five, and the average number of members in each household is only 3.[7] Hence, the size of five grids per household and a total of 37,070 (5 × 7,414) grids approximately give a reasonable space required for accommodating 7,414 households, being neither too tight nor too spacious. This setting roughly allows for a space of 1.66 grids per agent; the distance between two agents, in terms of social interactions, is frequently just one to two ‘clicks’ away. The space of 37,070 grids can be approximated by a rectangular space of 193×193 (= 37,249) grids. As far as the election event is concerned, the social interactions among them can be further restricted to voters only. The number of voters in the Taiwan 2012 Presidential election was 13,454,016. By scaling down this number equally by a factor of 1,000, we have a size of 13,454 agents. Hence, these 13,454 agents distributed in the rectangular space of 37,249 grids leads to a population density of 36.12% (=13,454/37,249).

3.2 The Tolerance and Exploration Capacities

Second, it is about the two capacities of each voter, *i.e.*, r and s. The setting of these two parameters is similar to that of Chie and Chen [2] except that each agent now has its own s (tolerance capacity) and r (exploration capacity), which are randomly determined from a uniform distribution $U[0.26, 0.75]$ (with an increment of 0.01) and $U[2, 6]$ (with an increment of 1). This heterogeneous design, therefore, has certainty equivalence to the homogenous case with $s = 0.50$ and $r = 4$. The settings of these two parameters are based on the following considerations.

The choice of the uniform distribution is motivated by its maximum entropy. Since the purpose of this study is to address the effect of the heterogeneity of agents on the accuracy of the prediction market, it will be useful to examine the case with the maximum degree of heterogeneity. Therefore, the uniform distribution is chosen for that purpose.

Table 1. Experimental parameters.

Parameter	Description	Value
N	Number of agents	13,454
d	Population density	36.12%
G	Grid size	193 × 193
s	Heterogeneous containing capacity	$U[0.26, 0.75]$
r	Heterogeneous radius	$U[2, 6]$
m	Number of candidates	3
v_1	Vote share of the Green candidate	[0.18, 0.47]
v_2	Vote share of the Blue candidate	$1 - v_1 - 0.03$
v_3	Vote share of the Orange candidate	0.03
R	Simulation runs	50

As to the ranges of the two uniform distributions, they are determined as follows. The tolerance capacity (s) is chosen between 0.26 and 0.75 for the convergence of the Schelling segregation dynamics. It is known that if the tolerance capacity is too low, it

[7] The official statistics (in Chinese) can be found at http://www.dgbas.gov.tw/public/ Attachment/111171361-171.pdf, p. 28, Table 23.

will become difficult for the model to converge. Therefore, we set 0.26 as the minimum, and, to be symmetric around 0.5, set 0.75 as the maximum.

The exploration capacity has a range from 2 to 6. This number determines the maximum number of neighbors that can reached by the agent. In terms of the von Neumann neighborhood, this number is 12 when $r = 2$, then increases to 28 ($r = 3$), 48 ($r = 4$), 80 ($r = 5$), 112 ($r = 6$), and 148 ($r = 7$). According to the social brain hypothesis [17, 18], humans' brain are capable of managing a maximum of just 150 friendships. By considering that social relations are not all political and not all friends are politically-oriented, we do not push the r to the maximum to 7, but stop at 6. The remaining capacity can, for example, be reserved for the social relations with people who are not much involved in elections, including non-qualified voters.

The two capacities introduced in Chie and Chen [2] may have good correspondence to some of the personality characteristics studied in psychology. The tolerance capacity seems to be related to a person's *agreeableness*, whereas the exploration capacity is a manifestation of a person's *extraverted* personality.[8] Of course, we understand that this correspondence may not be exact and should be rather suggestive. Nonetheless, it at least indicates the possibility that one can incorporate agents with different personalities into models of artificial agents and examine the effects of possible composition on the emergent aggregate outcomes, in our case, the performance of prediction markets. In addition, while Hayek did have a celebrated devotion to psychology [20], psychology plays almost no role, at least explicitly, in his 1945 magnum opus. Therefore, it remains of interest to know whether one should read between the lines to see if such connection may exist.

With this randomization of agents' capacities, we now have a society of heterogeneous agents with different degree of *agreeableness* (s) and *extraversion* (r), in addition to their political identifications. The simulation to be presented below is, therefore, a response to the formal inquiry into the role of these two personality characteristics in appraising the wisdom of crowds in the form of prediction markets.

3.3 Three-Party Politics

Third, it is about the number of political parties and candidates. Again, to mimic Taiwan's political reality, we consider an election involving three parties, two major ones (Green and Blue) accompanied by one marginal one (Orange). The size of the party is characterized by the actual share of the vote for each candidate, $\{v_i\}_{i=1}^{3}$. The vote share of the small party (the Orange Party), denoted by v_3, is fixed at 3%. The rest of the vote shares are determined as follows:

$$\{v_1, v_2\} = \{v_1, 1 - v_1 - 0.03\},\ v_1 = 0.18, 0.19, \ldots, 0.47,$$

[8] This correspondence can be justified as follows. Agreeableness is the degree to which a person needs pleasant and harmonious relations with others. Its manifestation facets include trust, straightforwardness, altruism, compliance, modesty, and tender-mindedness [19]. Therefore, it is assumed that agents with high agreeableness may find it easier to accommodate people with different political identifications, and hence have a higher tolerance capacity. Extraversion is the degree to which a person needs attention and social interaction. Its manifestation facets include warmth, gregariousness, assertiveness, activity, excitement, seeking, and positive emotions [19]. Hence, an agent with a higher degree of extraversion tends to have a greater ability to explore around and mingle with a larger number of neighbors. Therefore, it goes well with exploration capacity.

where v_1 is the vote share attributed to the Green Party, whereas v_2 is the one attributed to the Blue Party. We let v_1 to increase from 0.18 to 0.47 with an increment of 0.01. Hence, there are a total of 30 designs, and each is distinguished by a specific set of votes. We then generate 1,000 random samples from these 30 designs (on average, each design will be sampled 30 times).

Finally, for each sampled design, we have 50 runs, and each run is distinguish by a set of agents characterized by different values of s and r and political identities. To sum up, we have a total of 50,000 (1,000 × 50) runs.

4. SIMULATION RESULTS

The simulation results are organized into three parts. The first part concerns the forecasting accuracy of the prediction market. Since we have 30 designs ranging from very low vote shares (highly improbable events), such as v_3 = 0.03, to very high vote shares (highly probable events), such as v_1 = 0.79, we shall then first examine how the prediction market performs with respect to such a large variety of events (Section 4.1).[9] The second part concerns the forecasting accuracy comparison between the case where the two capacities of agents are heterogeneous and the case where they are homogeneous. In other words, we compare the performance of the heterogeneous design with that of the homogeneous design or the certainty-equivalent design (Section 4.2). The last part concerns the influence of the two capacities on individual earning performance (Section 4.3) and income distribution (Section 4.4).

To begin with, let us first make a note of the statistics to be employed below. Basically, statistics are the average taken over 50 runs. For the mean price, we first take the average of the price series for each run (Eq. (10)), and take the average of the average over these 50 runs (Eq. (9)).

$$\overline{p}_j = \frac{\sum_{l=1}^{50} \overline{p}_{j,l}}{50}, \quad j = 1, 2, 3, \quad l = 1, 2, ..., 50, \tag{9}$$

where

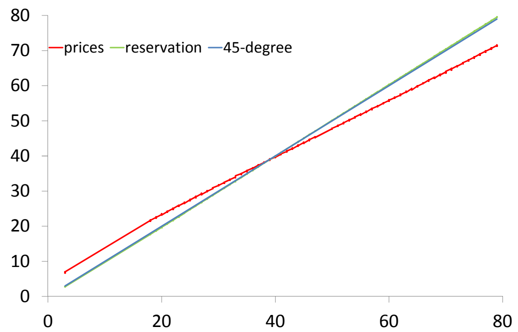
$$\overline{p}_{j,l} = \frac{\sum_{t_{j,l}=1}^{T_{j,l}} p_{j,l}(t_{j,l})}{T_{j,l}}, \quad j = 1, 2, 3, \quad l = 1, 2, ..., 50, \tag{10}$$

and $T_{j,l}$ are the transaction times of the futures j for the lth run.

4.1 Favorite-Longshot Bias

Fig. 6 first shows the forecasts made by the prediction market under different vote shares, called the price line (the red line). To see how well the forecast has been made, we also draw the 45-degree line (the blue line) as a reference. If the price line overlaps the 45-degree line entirely, then the prediction market has successfully forecast the vote share won by the candidate at each vote share. This is, however, not the case, as we have

9 For the general performance of the prediction markets, the interested reader is referred to [22, 23].



Note: The three lines above shows the correspondence between the true price (the actual vote share), on the x-axis, and the predicted price (the red line), the true price (the 45-degree blue line), and the reservation price (the green line) on the y-axis. The predicted price and the reservation price refer to the mean of the distribution. The means shown here are the average taken over all 50 runs. The line is drawn by using three segments: 3% is from Orange Party, followed by 18% to 47% from the Green Party, and 50% to 79% from the Blue Party. The rest of the segments are either absent or involve the use of interpolation.

Fig. 6. Favorite-longshot bias.

seen in Fig. 6. The two lines, the price line and the 45-degree line, intersect at 39%. Then there are positive deviations (overestimates) when the vote share is lower than 39% and negative deviations (underestimates) when vote share is higher than 39%, a typical result known as the *favorite-longshot bias* [21], to which we now turn.

The favorite-longshot bias is a phenomenon which was first found in the *pari-mutuel markets*, such as the horse racing market.[10] In pari-mutuel markets, the chances of uncertain events can be roughly characterized by *short odds* or *long odds*. Short odds, also known as *favorites*, refer to the bets on which the gambler will only receive a small return for his/her 'investment', which indicates that the chances of the betting events are good, whereas long odds, also known as *longshots*, refer to the bets on which the gambler will receive a good return on his/her 'investment', which indicates that the chances of the betting events are poor.

The first evidence on the favorite-longshot bias is documented in Griffith [24], which shows that horses with short odds yield on average higher returns than horses with long odds. This implies that, on average, compared to the empirical probabilities derived from real outcomes, the pari-mutuel market overpredicts the probabilities of long odds (longshots) and underpredicts the probabilities of short odds (favorites). If we plot the actual probabilities obtained from real outcomes on the x-axis and the market probabilities obtained from market prices on the y-axis, and also plot the 45-degree line as a reference, then the favorite-longshot bias implies that at the left end of the line, characterizing longshots, the market probabilities are greater than the empirical probabilities, but at the right end of the line, characterizing favorites, the market probabilities are smaller than the empirical probabilities. In other words, the slope of the market price line is less than 45 degrees. As shown in Fig. 6, by moving from left to right, the market price line crosses over the 45-degree line.

---

[10] In the pari-mutuel market, each gambler attributes a fund to a pool to participate in a bet on the outcome of an uncertain event, and the pool of funds will be distributed among all gamblers based on the outcome of the game.

Since Griffith [24], evidence of the favorite-longshot bias has accumulated in sport and betting markets and over different countries [25]. It is considered to be a very robust empirical finding in sport and betting markets [26, 27]. The pursuit for an understanding of its causes has drawn significant attention among economists due to its relation to the *efficient markets hypothesis* and even to the Hayek hypothesis [28]. Ottaviani and Sørensen [29], in their survey article, summarize seven possible explanations for the favorite-longshot bias. Among the seven, the one which is most close related to our study is the explanation proposed by Ali [30], which is an explanation based on the heterogeneous agents. A different version of the heterogeneous-agent approach focusing more on transaction and information costs was taken by [28].[11] One commonality of these heterogeneous-agent approaches is the use of the device of informed traders and uninformed traders. After a proper translation, this is essentially the same as our heterogeneous agents characterized by different tolerance capacities and exploration capacities. Hence, our finding can be considered to be that of the first simulation study which lends support to the heterogeneous-agent explanation for the favorite-longshot bias.

To show that the appearance of this bias is indeed very fundamental and is not caused by the sampling error, in Fig. 6, we also show the reservation price, as defined in Eq. (3) and exemplified in Fig. 5 (Panel C), by the green line. Now, we can see that this green line has a perfect overlap with the blue line. Therefore, the information distributed among the agents has a correct representation of the true price in the mean; the market, however, fails to aggregate this information, particularly for those highly likely or unlikely events.

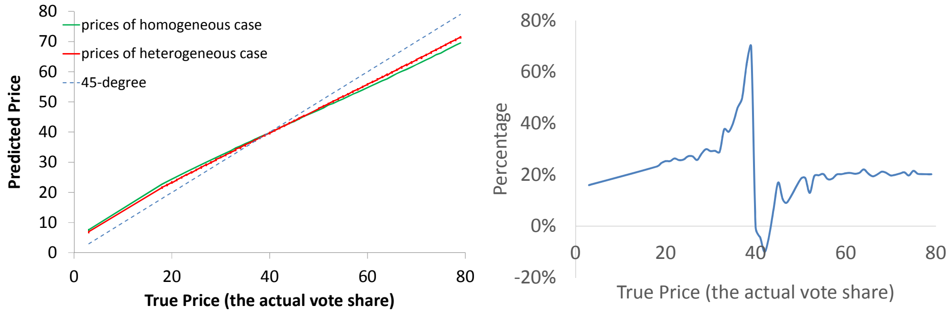
4.2 Heterogeneous vs. Homogeneous Designs

To see whether the prediction market composed of agents with a larger degree of heterogeneity can perform better than the market composed of homogeneous agents, Fig. 7 (the left panel) shows the price line coming out of the market with a homogeneous design (the green line) and, as a contrast, adds the red line which we have seen in Figs. 6 and 7 (the left panel). We can see that the favorite-longshot bias appears again with the green line (the homogeneous design), but it seems to be even a little outward-oriented as opposed to the red line (the heterogeneous design).

To make their relative positions easier to see, in the right panel of Fig. 7, we show the percentage of the change in the vertical distance between the 45-degree line and the red line (the heterogeneous design) as opposed to that of the green line (the homogeneous design), *i.e.*, the improvement rate in the vote share prediction error. The reduction rate beings at around 20% from both the left end (longshots) and the right end (favorites) and remains there for a long range until the vote share moves closer to the crossing point of these lines, *i.e.*, around 40%. There we can see that the percentage of improvement fluctuates quite substantially from a maximum of 70% at v = 39% to a minimum of −10% at v = 42%. In this crossing zone, both designs can have a quite good prediction of the actual vote share (the error rate is very low); hence given the small magnitude of the denominator, the improvement rate is naturally sensitive to a small change in the magni-

[11] Ottaviani and Sørensen [31] also provide an explanation for the favorite-longshot bias based on the heterogeneous beliefs of traders.

tude of the numerator. To sum up, the improvement rate is positive 93% of the time with a mean of 24.50% and a median of 20.72%. Nevertheless, the advantage of the heterogeneous design over the homogenous design is not statistically significant. In our case, since the uniform distribution has been applied to determine the agents' personality traits, the degree of heterogeneity is at its maximum. Therefore, this shows that the favorite-longshot bias is not very sensitive to the change in the degree of heterogeneity. Although a large degree of heterogeneity can improve the favorite-longshot bias, it can not do so to a degree to make it disappear.



Note: The left panel of the figure shows the predicted vote share (on the y-axis) along with the actual vote share (on the x-axis) under the homogeneous design (the green line) and under the heterogeneous design (the red line). To better observe the prediction accuracy, the 45-degree line (the blue dotted line) is also draw as a reference. The right panel then gives the percentage of the change in the error rate when the homogeneous design is replaced with the heterogeneous design.

Fig. 7. Favorite-longshot bias under homogeneous and heterogeneous capacities.

4.3 Individual Traits and Earning Performance

While the great heterogeneity of agents may not significantly beef up the market performance, it is interesting to see whether it has significance effects on individuals. Specifically, from a labor economics viewpoint, we are interested in knowing the role played by the two capacities, s and r, in the earning equation, *i.e.*,

$$\pi_i = \alpha + \beta_s s_i + \beta_r r_i + \varepsilon_i . \tag{11}$$

Regression (11) basically asks what is the expected trading profit of agent i conditional on his two capacities, s and r. To estimate (11), we generate a society of 13,454 agents and randomly determine their s and r using the distributions given in Table 1, *i.e.*, a triple $\{i, s_i, r_i\}$ (i = 1, 2, ..., 13454). This population of agents then trades in each of the 30 prediction markets (the 30 designs). In each market, their two capacities are fixed, but their political identity is randomly determined to be consistent with the individual design. With the given capacities and given political identities, we then run the same market (same design) for 50 times (50 different political configurations). We then observe their trading profit accumulated over these 30 designs and over these 50 runs, and the quadruple is $\{i, s_i, r_i, \pi_i\}$, where π_i is the accumulated profit. The regression (11) is then based on these 13,454 observations, which is large enough to allow us to have a good estimation of the general influence of s_i and r_i after leaving out the possible influence of political party (being smart party member or large party member) and location. The result is shown in Eq. (12).

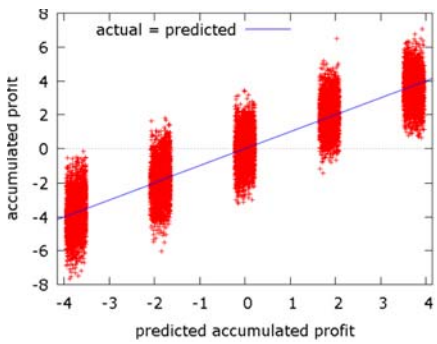
$$\pi_i = \underset{(0.0437)}{-7.90} + \underset{(0.0066)}{0.86}\, s_i + \underset{(0.0644)}{1.87}\, r_i + \varepsilon_i. \quad \overline{R}^2 = 0.86. \tag{12}$$

As usual, the numbers shown inside the parentheses are the standard errors of the corresponding estimates.

After controlling for some unnecessary complications (location and party), we can see that the earnings equation is essentially determined by the two capacities with an $\overline{R}^2$ (the adjusted coefficient of determination) of 0.86. In other words, 86% of the variation in trading profits can be explained by these two capacities. From Eq. (12), if we divide the estimate by its standard error, then we can have the t value, which is 284.6 for r and 13.39 for s. Obviously, both capacities are statistically significant, which suggests that agents with a higher tolerance capacity (agreeableness) and exploration capacity (extraversion) tend to have a higher earnings capacity when they engage in the prediction market.

Despite their statistical significance, in terms of the magnitudes of the two coefficients, Eq. (12) shows that the exploration capacity has a stronger effect on earnings than the tolerance capacity. This can be seen from the multiplication of the range of the parameter, r and s, and the respective estimates, 1.87 or 0.86. For r, the result is 7.48 (4 × 1.87), whereas for s, the result is 0.43 (0.5 × 0.86). The reason why the tolerance capacity has a mild effect is because, based on the Schelling migration process, it can only affect the choice of the neighbors in the physical space ($r = 1$), but not the social space ($r > 1$). When agents are heterogeneous in terms of their tolerance capacity, it will be harder to have a large cluster of agents with an identical political identity. Hence, an agent with a small s but a large r can still obtain access to a great diversity of agents that will make him/her become an informed trader. On the other hand, an agent with a large s may still be trapped in a small cluster; in this case, he/she can become uninformative if his/her r is small.

Fig. 8 gives the predicted accumulated profit, *i.e.*,



Note: The predicted profit is shown on the x-axis, and the actual profit is shown on the y-axis. The blue line in the middle is the 45-degree line. The profit earned by the agents is clearly segregated into five clusters, each corresponding to one of the five exploration capacities (r = 2, 3, 4, 5, or 6). The mean of a cluster has a tendency to move upward with r.

Fig. 8. Actual profit and predicted profit.

$$\hat{\pi}_i = -7.90 + 0.86 s_i + 1.87 r_i \tag{13}$$

on the x-axis and the actual accumulated profit, π_i on the y-axis. We can see that there are a large number of observations above or below $\hat{\pi}$ (the 45-degree line). These unfitted observations (the red points not on the 45-degree line) are expected since the operation of the double auction market has its own stochasticity and agents, by design, are zero-intelligent, indicating that some of their behavior is random.

Using Eq. (13), we can compare the earnings of the agents at the two extremes: one with $r = 2$, $s = 0.26$ and one with $r = 6$, $s = 0.75$. By Eq. (13), the former have an accumulated profit of –3.94, whereas the latter have an accumulated profit of 3.97. They are almost as symmetric as a pair of players in a zero-sum game. In finance, the former will be considered to be *uninformed traders*, and the latter will be considered to be *informed traders*. In our case, the former are the least informed traders, and the latter are the most informed traders. However, we provide a reason why the uninformed are uninformed and the informed are informed from a psychological perspective. Those agents who are the most uninformed and hence the poorest are so because they have an extremely low tolerance capacity (agreeableness) and a low exploration capacity (extraversion).

4.4 Income Distribution

In addition to knowing the effects of s and r on individual earnings, it will be also interesting to see their global impact on the society in terms of income distribution. As we mentioned earlier, the prediction market is a kind of financial market. The general issue that concerns economists is how investments made in these financial markets contribute to income inequality. Our simulation results may shed light on this issue.

To do this, we draw the Lorenz curve in Fig. 9 and also present the Gini coefficients for comparison purposes.[12] To see how the heterogeneity in the two capacities can contribute to the income inequality observed in the population of agents, the Lorenz curves are also draw separately, one being conditional on s only, and one being conditional on r only.

First of all, among the three coefficients, the Gini coefficient reaches its highest level when we restrict the heterogeneity of agents to their exploration capacity (r). The r-Lorenz curve (the purple one) tells us that this is so because the lowest 20% of the population, *i.e.*, the agents with an r of 0.2, has a share of 0% of the total income. In fact, their income is negative.[13] Notice that the agents in this group also differ in their tolerance capacities, with some featuring a large s, but, as Fig. 8 indicates, the negative effect caused by the low exploration capacity is so high that even after taking into account other advantages of the agents, such as their tolerance capacity and local diversity, their income is still negative (this can be seen in that no single point in the leftmost cluster in Fig. 8 has a positive earning).

[12] The x-axis of the Lorenz curve refers to the poorest x percent of the population, and the y-axis refers to y percent of the national income. Together, any single point (x, y) on the Lorenz curve indicates that the poorest x of the population shares y percent of the national income. If income is distributed equally among the whole population, then the Lorenz curve is a 45-degree line. On the other hand, any deviation from perfect income equality will cause the Lorenz curve to be convex, and the degree of convexity shows how unequal income is distributed among the population. Alternatively, it can also be measured by the Gini index, which indicates the percentage of the area between the 45-degree line and the Lorenz curve relative to the entire triangular area defined by the origin, (1,0) and (1,1). Obviously, this index is between 0 and 1; the higher the Gini index, the higher the degree of inequality.

[13] Hence, a proper transformation has been made before the usual Lorenz curve can be applied.

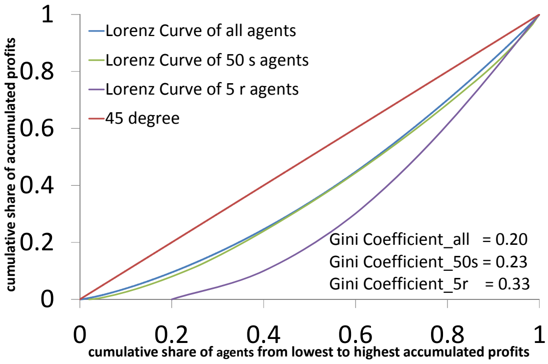


Fig. 9. Income distribution.

Second, compared to the exploration capacity, the effect of the tolerance capacity on income distribution is much milder. Its corresponding Lorenz curve, the *s*-Lorenz curve (the green one), has a Gini coefficient that is as low as 0.23, which is lower than the Gini coefficient of any human societies known to us.[14] Third, the unconditional Lorenz curve (the blue one) gives an overall effect with a Gini coefficient of 0.20. Hence, generally speaking, the prediction market studied by us, with the assumption of the two-dimensional personality trait, does not cause a severe income distribution problem in light of the empirical data. Although the exploration capacity may have the potential to threaten equality, other factors can mitigate its effect quite substantially. To sum up, the general distinction between informed and uninformed traders, a subject well studied in finance, does not contribute substantially to the income distribution inequality in our prediction market.

5. CONCLUDING REMARKS

In this article, we simulate an agent-based prediction market to learn about the information aggregation through the market mechanism. Specifically, we examine the role of the heterogeneity of agents in this mechanism. At the macro level, we find that agents who are heterogeneous in the two capacities, tolerance capacity (agreeableness) and exploration capacity (extraversion), can make the mechanism perform better, while the observed improvement is rather limited. At the individual level, we find both of the capacities can contribute to the earning performance.

For the latter, this is the first study showing that agents with different personality traits may have different trading performance. To what extent this finding is empirically relevant is an issue deserving further investigation through human-subject experiments or through empirical studies on the prediction market data. In fact, if personality traits have an effect on the earning performance, then by market selection it will be interesting to know whether the market participants of prediction markets, as a whole, generally have different personalities as opposed to non-market participants. In addition, the personality traits can have a determining role because in our model information becomes available only through personal contact and it is this assumption that makes the significance of personality traits self-evident. In the current digital world, social media have

[14] Based on World Bank data, Romania in 2010 actually had a Gini index of 0.24, which is the one closest to ours.

formed various social networks by which personal contacts are less direct, which may also downplay the role of personality traits in obtaining information. Hence, whether the role of personality traits in the distribution of information among individuals can be fundamentally different from a physical network to a social network is a direction for further studies.[15]

Although our agents are heterogeneous in the two capacities, they are all zero-intelligent agents and they are all limited to one entrance to the market. The zero-intelligent agent has no memory and is not able to learn. This design can be inconsistent with the assumed personality, such as extraversion. Hence, in the future, trading behavior with various learning behaviors should also be considered to be part of personality. Chen [2] has associated different kinds of learning models with different cognitive capacities; in this light, different learning models associated with different personalities may also be considered in the future. In this case, our agents are not only heterogeneous in terms of personality, but heterogeneous in terms of learning behavior. We can initiate this line of research by first starting with different kinds of learning behavior, for example, by replacing zero-intelligence agents with reinforcement-learning learning agents as a starting point.

Of course, learning can make sense if agents are allowed to enter the market repeatedly. It is, therefore, desirable to extend the current sampling scheme without replacement into the one with replacement. In the beginning, agents have an equal probability of being sampled; then in the future even this probability can vary with agents; hence, the trading aggressiveness of agents is also heterogeneous. With these aforementioned possible research directions, this paper can then be regarded as the first step toward the understanding of the heterogeneities of agents and their role in information aggregation. As a continuation of Hayek [1], in addition to the market as an information aggregation mechanism, we are also interested in the market as an aggregated personality, and it is through this aggregated personality that the information is aggregated.

REFERENCES

1. F. A. Hayek, "The uses of knowledge in society," *American Economic Review*, Vol. 35, 1945, pp. 519-530.
2. B.-T. Chie and S.-H. Chen, "Spatial modeling of agent-based prediction markets: Role of individuals," in *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, 2014, to appear.
3. V. Smith, "Markets as economizers of information: Experimental examination of the 'Hayek Hypothesis'," *Economic Inquiry*, Vol. 20, 1982, pp. 165-179.
4. T. Schelling, "Dynamic models of segregation," *Journal of Mathematical Sociology*, Vol. 1, 1971, pp. 143-186.
5. N. Vriend, "Was Hayek an ACE?" *Southern Economic Journal*, Vol. 68, 2002, pp. 811-840.
6. P. J. Boettke, ed., *Socialism or the Market: The Socialist Calculation Debate Revisited*, Vol. 9, Routledge, UK, 2000.
7. A. Keen, *The Cult of the Amateur: How Today's Internet Is Killing Our Culture*, Doubleday, UK, 2007.

[15] Actually, Qiu, Rui, and Whinston [33] have studied the performance of the prediction market involving social media, such as Twitter. Social media can impact the formation of social networks and the information flow. It remains a challenging issue to model the prediction markets in which different events may be embedded with different social media and different social networks.

8. M. Bauerlein and S. Walesh, "The dumbest generation: How the digital age stupefies young Americans and jeopardizes our future," *Leadership and Management in Engineering*, 2009, Vol. 9, p. 100.
9. J. Tirole, "On the possibility of trade under rational expectations," *Econometrica*, Vol. 50, 1982, pp. 1163-1182.
10. T. Yu and S.-H. Chen, "Agent-based model of the political election prediction market," in *Proceedings on the 12th International Workshop on Multi-Agent-Based Simulation*, 2011, pp. 117-128.
11. A. Othman, "Zero-intelligence agents in prediction markets," in *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, 2008, pp. 879-886.
12. D. Gode and S. Sunder, "Allocative efficiency of markets with zero intelligence traders: Market as a partial substitute for individual rationality," *Journal of Political Economy*, Vol. 101, 1993, pp. 119-137.
13. S.-H. Chen, "Varieties of agents in agent-based computational economics: A historical and an interdisciplinary perspective," *Journal of Economic Dynamics and Control*, Vol. 36, 2012, pp. 1-25.
14. R. Axelrod, "The dissemination of culture: A model with local convergence and global polarization," *Journal of Conflict Resolution*, Vol. 41, 1997, pp. 203-226.
15. D. Centola, J. Gonzalez-Avella, V. Eguiluz, and M. S. Miguel, "Homophily, cultural drift, and the co-evolution of cultural groups," *Journal of Conflict Resolution*, Vol. 51, 2007, pp. 905-929.
16. S.-H. Chen, C.-L. Chang, and M.-C. Wen, "Social networks and macroeconomic stability," *Economics: The Open-Access*, *Open-Assessment E-Journal*, Vol. 8, 2014, http://dx.doi.or/10.5018/economics-ejournal.ja.2014-16.
17. R. Dunbar, "The social brain hypothesis," *Evolutionary Anthropology*, Vol. 6, 1998, pp. 178-190.
18. R. Dunbar and S. Shultz, "Evolution in the social brain," *Science*, Vol. 317, 2007, pp. 1344-1347.
19. L. Borghans, A. Duckworth, J. Heckman, and B. Weel, "The economics and psychology of personality traits," *Journal of Human Resources*, Vol. 43, 2008, pp. 972-1059.
20. F. A. Hayek, *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*, The University of Chicago Press, IL, 1952.
21. E. Snowberg and J. Wolfers, "Examining explanations of a market anomaly: Preferences or perceptions?," in D. B. Hausch and W. T. Ziemba, eds., *Handbooks in Finance: Handbook of Sports and Lottery Markets*, Elsevier, North-Holland Press, The Netherlands, 2008, pp. 103-136.
22. J. Berg, F. Nelson, and T. Rietz, "Prediction market accuracy in the long run," *International Journal of Forecasting*, Vol. 24, 2008, pp. 285-300.
23. S.-H. Chen, C.-Y. Tung, C.-C. Tai, B.-T. Chie, T.-C. Chou, and S.-G. Wang, "Prediction markets: A study on the Taiwan experience," in L. Williams, ed., *Prediction Markets: Theory and Application*, Routledge, Chapter 11, pp. 137-156.
24. R. Griffith, "Odds adjustment by American horse-race bettors," *American Journal of Psychology*, Vol. 62, 1949, pp. 290-294.
25. W. T. Ziemba, "Efficiency of racing, sports, and lottery betting markets," in D. B.

Hausch and W. T. Ziemba, eds., *Handbook of Finance: Handbook of Sports and Lottery Markets*, North Holland Press, 2008, pp. 183-222.
26. R. Thaler and W. Ziemba, "Anomalies: Parimutuel betting markets: Racetracks and lotteries," *Journal of Economic Perspectives*, Vol. 2, 1988, pp. 161-174.
27. E. Feess, H. Müller, and C. Schumacher, "The favorite-longshot bias and the impact of experience," *Business Research*, 2014, pp. 1-18.
28. W. Hurley and L. McDonough, "A note on the Hayek hypothesis and the favorite-ongshot bias in parimutuel betting," *American Economic Review*, Vol. 85, 1995, pp. 949-955.
29. M. Ottaviani and P. N. Sørensen, "The favorite-longshot bias: An overview of the main explanations," in D. B. Hausch, W. T. Ziemba, eds., *Handbook of Sports and Lottery Markets*, North-Holland/Elsevier, 2008, pp. 83-101.
30. M. Ali, "Probability and utility estimates for racetrack bettors," *Journal of Political Economy*, Vol. 85, 1977, pp. 803-815.
31. M. Ottaviani and P. N. Sørensen, "Aggregation of information and beliefs in prediction markets," Finance Research Unit, Department of Economics, University of Copenhagen, 2007.
32. S.-H. Chen, "Reasoning-based artificial agents in agent-based computational economics," in K. Nakamatsu, L. Jain, eds., *Handbook on Reasoning-based Intelligent Systems*, World Scientific, Singapore, 2013, pp. 575-602.
33. L. Qiu, H. Rui, and A. Whinston, "A twitter-based prediction market: Social network approach," SSRN: http://dx.doi.org/10.2139/ssrn.2047846.

Bin-Tzong Chie (池秉聰) is an Assistant Professor of the Department of Industrial Economics at Tamkang University and a Research Fellow of AI-ECON Research Center at National Chengchi University in Taiwan. He holds a Ph.D. in Economics from the National Chengchi University. His research interest includes genetic programming, agent-based computational economics, and experimenttal economics.

Shu-Heng Chen (陳樹衡) is a Distinguished Professor in the Department of Economics at National Chengchi University, where he also serves as the Director of the AI-ECON Research Center as well as the Experimental Economics Laboratory. He is the editors-in-chief and the editor for a number of journals, such as Journal of New Mathematics and Natural Computation, Journal of Economic Behavior and Organization, and Journal of Economic Interaction and Coordination. Dr. Chen received a Ph.D. in Economics from the University of California at Los Angeles in 1992. His major research interests include agent-based computational economics, experimental economics, behavioral economics, and artificial intelligence.