Efficiency of the Experimental Prediction Market: Public Information, Belief Evolution, and Personality Traits

Bin-Tzong Chie[1] and Chih-Hwa Yang[2]

Abstract

This paper examines the ability of markets to aggregate information so that the price generated from the market contains the best estimate of all the available information. The paper investigates how individuals "update" their initial beliefs from their public and private information in light of market prices. In particular, the paper looks at individuals' weighting of public information versus private information. Also, the volume of information in the market via an increased number of traders with private information has a positive impact on the quality of the market price. Lastly, the personality traits of the traders seem to provide some positive impact if the traders are diverse in terms of the proportion of "efficient and organized" traders in the market.

JEL classification numbers: D82, C91, C92

Keywords: Experimental economics; Belief; Prediction markets; Market efficiency; Personality traits

[1] Professor, Department of Industrial Economics, Tamkang University, Taiwan.
[2] Ph.D program in Industrial Economics, Tamkang University, Taiwan.

1 Introduction to the prediction market

Economists have tried to identify possible causes of market efficiency. The prediction market applies a trading mechanism to disclose and turn fragmented information into a collective market price. George Neumann, Robert Forsythe and Forrest Nelson started to utilize market mechanism to predict the election outcomes in 1988. They established the Iowa political stock market that allowed faculty, staff and students to trade at the University of Iowa. What was being traded at that time were the futures contracts of voting shares for George H. W. Bush, Michael Dukakis, and the other presidential candidates.[1] The results had shown that the prediction market was rather accurate, outperforming concurrent polls (Hahn and Tetlock 2006). In 1992, the Iowa political stock market was renamed the Iowa Electronic Market (IEM) and traded a variety of objects with the mechanism of futures market. As the networking infrastructure became advanced, the IEM participants could come from all over the world.

The efficiency of the prediction market has been supported empirically and theoretically by Wolfers and Zitzewitz (2004) and Gjerstad (2005). Even a small scale

[1] The investment limit of a trader was from $5 to $500 dollars for a futures contract during the period from June 1, 1988 to the morning of the presidential election day. It was a zero-sum game with participants transacting spontaneously based on their own beliefs (preferences). At the end of the day, all the contracts were settled in accordance with the election results.

market can be efficient (Christiansen, 2007). Researchers have endeavored to unravel information efficiency embedded in the market. For example, Forsythe et al. (1992) and Forsythe, Rietz and Ross (1999) successfully predicted the outcome of the US presidential election with the political stock market.

However, it remains questionable whether the market can always aggregate information efficiently. Before the result of the U.S. presidential election of 2016, the IEM still predicted that Hillary Clinton will beat Donald Trump.[2] Haruvy, Lahav and Noussair (2007) found that market individuals' beliefs about prices are informative for researchers to infer the movements of future market price. While remaining strong beliefs, the IEM should nevertheless offer a false prediction. When the traders had gradually realized the electoral outcome Hillary's contract began to crash and Trump's contract suddenly became very hot.[3]

On the other hand, Ottaviani and Sørensen (2007) suggested that under certain circumstances it might be difficult for the market to achieve efficiency, such as no trade theory, herd behavior and manipulation. In this paper, our goal is to apply an experimental prediction market to revisit market efficiency and explore the influence

[2] http://www.cbsnews.com/news/markets-predict-hillary-clinton-will-beat-donald-trump/
[3] For relevant information, please refer to the link https://iemweb.biz.uiowa.edu/graphs/graph_PRES16_VS.cfm

of individual beliefs and personality traits under an experimental market.

Through market mechanism, many researchers indicate that the advantage of markets. Hanson (2006) has introduced many predictive phenomena that have long existed in markets. For example, orange-juice futures can help the National Weather Service improve weather forecasts (Roll 1984); horse race betting beats experts' forecasts (Figlewski 1979); stock markets reacted earlier than the NASA investigation panel in pointing to the company responsible for the Challenger accident (Maloney and Mulherin 2003). These findings support the functionality of information aggregation of markets and the importance of information disclosure from market participants.

The rest of the paper is organized as follows. Section 2 reviews the literature on market efficiency and experimental markets. Sections 3 and 4 detail our experimental design, data analysis. Section 5 presents the results and a discussion, and is followed by the conclusion in Section 6.

2 Literature review

2.1 Prediction markets

Prediction markets apply the market mechanism to achieve the information aggregation. With participants motivated by profitable opportunity, the prediction market not only provides up-to-date information, but also saves tremendous cost and time compared to polls. In other words, by properly aligning the incentives of participants, the prediction market can reveal true preferences at the aggregate level economically. Berg, Nelson, and Rietz (2008) demonstrated that the IEM outperformed polls for 76 percent of the elections between 1988 and 2004. Snowberg, Wolfers, and Zitzewitz (2007) studied a longer period of election samples and found a stable pattern of electoral outcomes affecting financial markets. These findings suggest that markets would try to account for valuable information by all means. The information advantage has been studied through the data of the IEM (Berg et al., 2008). In particular, Berg et al. (2008) found that some traders might have more information than others, being defined as “marginal traders”. As the marginal traders search for profitable opportunities, they improve the market efficiency.

Many studies have tried to identify possible factors of market efficiency. Figlewski (1978) showed that it is mathematically evident that the more homogeneously distributed information is, the more efficient a prediction market can be. Oliven and Rietz (2004) analyzed the questionnaire data of the IEM traders and found that despite of participants’ anomalous and irrational behavior, the prediction market can be as efficient. These findings corroborate the Hayek hypothesis. According to Hayek (1945), although bearing bias, the market is able to aggregate individual opinion and form a “price” that not only conveys information but also directs resource allocation.

On the contrary, there are findings for market inefficiency (Rosenberg, Reid and Lanstein 1985; Hazen 1987). One of the noticeable facts on market inefficiency is the systematic pricing bias, which is called the *favorite-longshot bias* (FLB). This bias is attributed to the traders’ tendency in overestimating low probability events or underestimating high probability events (Griffith, 1949). Isaacs (1953) demonstrated that the bias resulted from the profit maximization behavior of the monopoly informed trader. Weitzman (1965) discussed the willingness of risk-seeking traders to accept a lower expected payoff, causing over investment in low probability events. Heterogeneous beliefs among traders might also contribute to the bias (Ali 1977).

Ottaviani and Sørensen (2010) developed a series of testable features, including the quality of private information, the number of traders, the number of outcomes, common unpredictable errors, the level of participation, and the prior distribution of beliefs. These have provided us with clues for constructing the experimental market.

2.2 Experimental markets

The investigation of the efficient markets hypothesis through experiments can be dated back to the 1940s (Chamberlin 1948). Chamberlin allowed students to bargain with each other one by one, rewarding the advantaged with money. Nonetheless, prices in his experimental market failed to converge to equilibrium status.

Vernon Smith indicated that the failure could be a result of the unavailability of *public information* for participants to learn and coordinate, which was later rectified in the experiments of Smith (1962, 1982). Based on public market price information, Forsythe and Lundholm (1990) found evidence to support the view that trading experience and a common knowledge of dividends would also contribute to the discovery of a rational expectations-based equilibrium price.

Plott and Sunder (1982) discussed whether individual participants could recognize assets with the best payout likelihood in a hypothetical stock market. Participants were asked to trade with three Arrow-Debreu securities based on partial information. In their study, Plott and Sunder (1988) found that if the information is homogeneously distributed or complete, rational expectations would be the best predictive model. The results show that the importance of information distribution matters.

Several studies have investigated *information heterogeneity*. Copeland and Friedman (1987) found that wide bid-ask spreads could be caused by traders' information heterogeneity. Plott and Sunder (1988) stated that the double auction market is an efficient mechanism for aggregating information and leads to rational expectations being in equilibrium. Ackert, Church and Zhang (2002) conducted experiments in a market consisting of both well-informed and less-informed traders, and they found that the former would take advantage of their information and outperform the latter.

Plott, Wit and Yang (2003) performed betting market experiments, in which they found *private information* could better explain traders' behavior than aggregated *public information* (price). This explains why most traders might feel contradictory and struggle between *public* and *private* information. As traders' experiences accumulate, Plott et al. (2003) found that the market would converge to rational expectations-like performance. Their results confirm that information aggregated at the market level would be influenced by time and the traders' experience.

There are studies that investigate manipulators in experimental markets. Deck et al. (2013) showed that market participants can be distinguished by the level of information possessed. Informed traders (the manipulators) have complete information and often act in a disguised manner in an attempt to dictate the market. Ordinary traders are those who own partial information and adjust views based on observed market prices. In addition to manipulators and ordinary traders, there are forecasters who have no information at all and are often beguiled by the behavior of manipulators. Deck et al. (2013) have shown that the interaction between manipulators and forecasters results in market failure.

As shown in the above literature discussion, there exists no unanimous agreement on market efficiency. In this paper, we attempt to control market variables by setting up an experimental prediction market in accordance with Deck, Lin and Porter (2013). In our small-scale experimental prediction markets, we test market efficiency by altering factors such as the distribution of private information, initial beliefs, public information, the number of traders, the evolution of beliefs, and the heterogeneity of personality traits. This study will not focus on the effects of manipulators. Our experimental market consists of ordinary traders only. Through the double auction mechanism, we intend to observe market performance without fully informed traders so as to investigate other possible endogenous causes for efficiency or inefficiency.

3 Experimental design

This section will first introduce our subjects, experimental design followed by a description of the distribution of information and beliefs.[4]

3.1.1 Subjects

We conducted 15 experimental sessions and had totally recruited 183 subjects.[5] The

[4] Appendix A summarizes the details of the experimental design and session description in Tables A1 and A2, respectively.

[5] The experiments were constructed using z-trees (Fischbacher 2007) and conducted in the

subjects were college or graduate students recruited from universities in Taiwan. Each subject only attended one of our experimental sessions. Repeat participants were not allowed. The number of subjects (m) per experiment session ranged from 5 to 19. Each session lasted for 2.5 hours, including 1 hour of introduction, 1 hour of trading, and 0.5 hours of personality testing. The average pay rate for our experiment is about NT$160 per hour. The total actual reward depends on performance, ranging from NT$260 to NT$500, with the average being NT$400 (2.5 × 160).

3.1.2 Information distribution design

Through the pre-experiment introduction, every subject was made aware of the fact that information is valuable. In our experiment, the information distribution of each session was represented by a set of N balls in a black box, which were k yellow balls and ($N - k$) white balls. Each subject was asked to draw n balls with replacement from the black box and keep the information privately. The drew n balls were x yellow balls and ($n - x$) white balls. After all subjects' drawing, the experimenter then drew a ball from the black box and announced the color to all subjects. The ball was represented by y, which was 1 if the ball was yellow, otherwise, y was 0. After that, subjects were instructed to predict the percentage of yellow balls. For each participant, according to each $p = \frac{k}{N}$,

Experimental Economics Lab at National Chengchi University. Further information about the institution can be found at http://futures.nccu.edu.tw/.

the probability of observed x (private information) plus y (public information) yellow balls can be represented by the following probability mass function:

$$f(x, y; n, p) = \Pr(X = x + y) = \binom{n}{x} p^{x+y}(1-p)^{n+1-x-y},$$

$$x = \{0, 1, \ldots, n\}\ ; y = \{0,1\}. \tag{1}$$

In our experiment, $N = 10$, $n = 2$, $k = 1, 2, \ldots, 9$.

3.1.3 Beliefs of subjects

After observing the private and public information, all subjects were requested to form and report their *beliefs*. For example, if a subject observed two yellow balls and the public information was also a yellow one, he or she might think that over 80 percent of the balls in the black box were yellow. If the market price fell below 80, this participant would go long. For our analysis, each subject had to answer the two questions below:

1. How many yellow balls did you observe, including your draw and the public announcement?
2. What do you believe is the percentage of yellow balls contained in the black box?

Trading was launched after all subjects completed above two questions.

According to the probability mass function (1), suppose the prior probabilities of k = 1, 2, 3, …, 9 balls were uniform. We can infer subjects' posterior beliefs predicted by the following Bayesian updating.

$$\mathrm{P_r}(k = i|X = x + y) = \frac{\mathrm{P_r}(X|k = i)\mathrm{P_r}(k = i)}{\mathrm{P_r}(X)} = \frac{\mathrm{P_r}(X|k = i)\mathrm{P_r}(k = i)}{\sum_{j=1}^{9} \mathrm{P_r}(X|k = j)\mathrm{P_r}(k = j)},$$

$$\text{where } i = 1, 2, \ldots, 9.$$

Then the expected posterior beliefs for the number of yellow balls can be inferred by $E(k|X = x + y) = \sum_{j=1}^{9}(k = j|X = x + y)\, k$. The Bayesian percentage of yellow balls contained will be 10 × $E(k|X = x + y)$.

3.1.4 Trading

Each session had two treatments: one was based on the ratio of yellow balls (i.e., the *ratio treatment*) and the other on whether yellow balls outnumbered white ones (i.e., the *winner-take-all treatment*). The experimenter should announce which treatment he will begin with. Each treatment lasted for four rounds and each round took three minutes. Based on the true value K (= 10×k), the market price was defined, lying between 0 (i.e., no yellow ball) and 100 (i.e., all yellow balls).

Each round, traders will be lending 1,000 tokens from the experimenter. Traders are allowed to buy or sell with these tokens. At the end of the market round, the 1,000 tokens will be returned to the experimenter. In this study, the trading mechanism of the futures market is applied. Subjects' contract holding position can be positive or negative. Positive means net buying, otherwise means net selling. During the trading period, the subjects' token holding is just temporary, for that all contract holding will also be settled with token at the end of each round. The goal of each subject is try to use all available information to trade. The final payoff is a linear transformation of the tokens subjects earn.

A standard continues double auction trading mechanism is adopted in this experimental market. For both treatments, the market reference price started at 50, representing maximum uncertainty of the percentage of yellow balls. Under the double auction mechanism, subjects could place their bid or ask prices by integer ranging from 0 to 100, or accept what was offered by other traders. After subjects placed their bid or ask prices, there were several bid and ask prices; they were maintained in the bid and ask queues, ordered first by offer price and then by time of placement.

The maximum possible loss will be deducted first from the token account for each traded contract. For example, If a contract price of a transaction is 60, the buyer pays 60, and the seller pays 40 as a deposit. If the final settlement price is 80, the buyer will receive 80 and the seller will only receive 20 tokens back. If the subject does not have enough tokens, no further orders can be placed.

At the end of the trading, subject *i*'s post-transaction belief (b_i^p), which would be associated with additional reward or punishment, was elicited. The incentive was equal to $50 - |b_i^p - K|$ in token value. This settlement process took place after all subjects declared their post-transaction beliefs. Each round of the market clearing was a zero-sum game.

3.1.5 Settlement

The experimental prediction market applied a cash settlement. For the settlement price (S) and holding position (h_i), trader *i* was to receive ($S \times h_i$) in token value. Here, S was dependent on the type of the two treatments applied. For the *winner-take-all* (WTA) *treatment*, the settlement price S was determined as follows:

$$S = \begin{cases} 100, & \text{if } K > 50; \\ 50, & \text{if } K = 50; \\ 0, & \text{if } K < 50. \end{cases}$$

For the *ratio treatment*, the settlement price S was defined as $S = K$.

4 Experimental data analysis

A total of 15 sessions were conducted; each session had 8 rounds. In this study, the characteristic features of each round (*r*) was calculated as a data record. Therefore, the sample size of the experimental data was 120.

We tried to measure market efficiency by means of the following procedures. The purpose of this experiment is to compare how the market prices compare to predicted market prices, presumably based on the objective probabilities of the number of yellow balls available.

We first define the objective probabilities that prices should be compared to. For example, if there are 7 subjects and they observe 9 yellow balls and 6 white balls, what is the probability that there were $k = 1, 2, 3, \ldots, 9$ yellow balls in the black box? The answer to that question is determined by Bayes' theorem. Then the probability that draws are made from a black box with k yellow balls given the sample

$$\mathrm{P_r}(k = i|p) = \frac{\mathrm{P_r}(p|k = i)\mathrm{P_r}(k = i)}{\mathrm{P_r}(p)} = \frac{\mathrm{P_r}(p|k = i)\mathrm{P_r}(k = i)}{\sum_{j=1}^{9} \mathrm{P_r}(p|k = j)\mathrm{P_r}(k = j)},$$

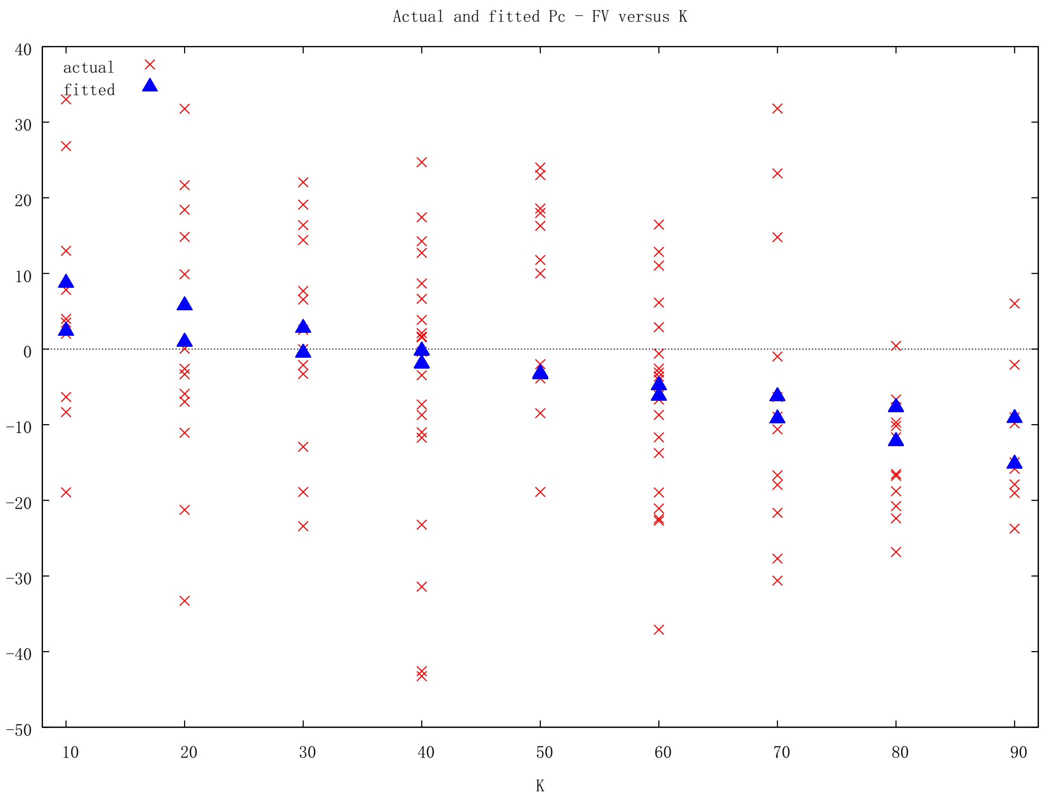
$$\text{where } i = 1, 2, \ldots, 9 \text{ and } p = \frac{k}{10}.$$

For each $k = j$, where $j = 1, 2, \ldots, 9$, the probability of each of these samples is binomial distribution $X \sim \mathrm{b}(n, p)$ with $n = 15$, $p = j / 10$, and d equal to the number of yellow balls drawn in the sample. The probability of drawing d yellow balls in a sample of size n conditional on k yellow balls in the black box is $f(d,n,p) = \mathrm{P_r}(X = d) = \binom{n}{d}(p)^d(1-p)^{(n-d)}$.

In the example we took this as $d = 9$. Suppose the prior probabilities of $k = 1, 2, 3, \ldots, 9$ balls were uniform. All priors are 1/9. So in the example, the probabilities are (0.0000, 0.0011, 0.0185, 0.0979, 0.2444, 0.3306, 0.2356, 0.0688, 0.0031). In the example with group size seven the expected value of the number of balls in the black box is 0.5882 when 9 of 15 balls drawn are yellow, and that is true regardless of the number of yellow balls in the black box. Suppose there are five yellow balls in the black box but 9 of the 15 draws are yellow. Then the measure that the we use as the *fundamental value* is what the subjects could reasonably infer given the information available to them.

Then, we used closing price (p_c) of each round as a proxy of market performance.[6] We define the *fundamental-based pricing error* as p_c minus the respective fundamental value. Figure 1 shows the *pricing error* ($p_c - FV$) (y-axis) versus each true value (K) (x-axis) for both treatments. The vertical bars mark the range of the mean plus and minus one standard deviation.

6 In fact, Deck et al. (2013) justify using closing prices (in addition to mean prices) by stating, "previous studies have found that closing prices typically better reflect aggregated information."



[Figure 1 about here]

4.1 Market efficiency

From Figure 1, we observe diverging slopes between the two treatments. The following regression will be applied to test whether the coefficients of treatment dummy and designed true value (K) are significant or not.

$$Error = \beta_0 + \beta_1 D_t + \beta_2 K + \beta_3 D_t K + \varepsilon , \qquad (3)$$

where ε is assumed to be an i.i.d. random variable that is normally distributed and D_t is the treatment dummy ($D_t = 1$ for the *WTA* treatment, and 0 for the *ratio* treatment).

Below is the regression result of our experimental data.

$$Error = \underset{(0.000)}{26.95} - \underset{(0.092)}{8.99} D_t - \underset{(0.000)}{0.592} K + \underset{(0.051)}{0.19} D_t K, \quad \overline{R}^2 = 0.47,$$

$$Error = \underbrace{11.751}_{(0.0263)} - \underbrace{7.912}_{(0.2571)} D_t - \underbrace{0.299}_{(0.0007)} K + \underbrace{0.155}_{(0.2102)} D_t K, \quad \overline{R}^2 = 0.09,$$

where the *p*-values are presented in the parentheses and $\overline{R}^2$ is the adjusted R^2. For the ratio treatment ($D_t = 0$), the constant term (11.751) shows that the market price tends to overprice the fundamental value when K is small. As K increases, *Error* will be alleviated until K reaches a critical value, which is around 39.3 (=11.751/0.299). When K is larger than 39.3, the market will tend to underprice the fundamental value. The WTA treatment also suffers, but non-significant with a milder constant term 3.839 (= 11.751 – 7.912) and slope -0.144 (= -0.299 + 0.155). Although this pattern may compliant with the FLB, the treatment and interaction dummies are not significant. We simply apply regression (4) to test whether the coefficient of designed true value is significant or not.

$$Error = \beta_0 + \beta_1 K + \varepsilon, \tag{4}$$

The result is presented below

$$Error = \underbrace{7.589}_{(0.1029)} - \underbrace{0.219}_{(0.0019)} K, \quad \overline{R}^2 = 0.09.$$

Through the low $\overline{R}^2$, coefficient of determination, shows less than 10% of the variance in the error that is explained by the designed true value. However, the perception of participants may also be subject to their own beliefs.

For the error measurement of the aforementioned influence from perception, we define the following variables:

- Mean initial belief: $\bar{b} = \sum_{i=1}^{m} \frac{b_i}{m}$, where b_i is the initial belief of subject i; and
- Mean post-transaction belief: $\bar{b}^p = \sum_{i=1}^{m} \frac{b_i^p}{m}$, where b_i^p is the belief reported by subject i at the end of transaction.

The null hypothesis we proposed here is that the pricing error is a pure white noise, which is

$$H_0: \beta_0 = \beta_1 = 0$$

Table 1: Market efficiency test

Error Measurement (dependent variable)		$\hat{\beta}_0$	$\hat{\beta}_1$	95% CI of $\hat{\beta}_0$	95% CI of $\hat{\beta}_1$	$\overline{R}^2$
(a)	$p_c - FV$	7.589	-0.219***	-1.74, 16.92	-0.34, -0.10	0.09
		(0.1029)	(0.0019)			
(b)	$p_c - \bar{b}$	3.221	-0.1261*	-5.14, 11.58	-0.25, -0.00	0.05
		(0.4225)	(0.0488)			
(c)	$p_c - \bar{b}^p$	2.414	-0.092*	-3.28, 8.11	-0.19, 0.01	0.05
		(0.3787)	(0.0678)			
(d)	$\bar{b} - FV$	4.368*	-0.093**	-0.91, 9.65	-0.18, -0.00	0.05
		(0.0978)	(0.0452)			

(e)	$\bar{b}^p - FV$	5.175* (0.0936)	-0.127** (0.0105)	-0.10, 11.34	-0.22, -0.03	0.06

Note: *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively. p-values in parentheses.

The test results are shown in Table 1. The error measurements of (a) to (c) failed to reject $\beta_0 = 0$. However, if we look at β_1, the slope of the regression line, it can also be interpreted as a proxy of the degree of market efficiency. The smaller $\hat{\beta}_2$ is in absolute value terms, the more efficient the market is. According to Table 1, the market price is more efficient in reflecting the mean post-transaction belief ($\bar{b}^p$) than the fundamental value (*K*).

From (5), we investigate $\bar{o} - K$, in which all the coefficients are unable to reject H_0. Under these circumstances, the mean observation ($\bar{o}$) is an unbiased estimate of the fundamental value. Then, what causes the market price to deviate from its fundamental value? If we further take look at (6) and (7), we will have $\bar{b} - K$ and $\bar{b}^p - K$. Except for $\hat{\beta}_1$, all coefficients are significantly different from zero. This means that belief could be the reason to cause the market price deviate from its fundamental value.

This might be the case where the majority of subjects did not take the sampled observations as their beliefs. They tended to attach too much weight to the probability of the lower observations and too less weight to the higher ones. This phenomenon can be captured by the probability weighting function in Tversky and Kahneman's (1992) study and followed by Camerer and Ho (1994), Tversky and Fox (1995), Wu and Gonzalez (1996), Abdellaoui (2000), Abdellaoui, Vossmann and Weber (2005), and Stott (2006).

However, the probability weighting function may not be the only reason for the market efficiency distortion. The bargaining process might also play a role (Knetsch, Tang and Thaler 2001; Loomes, Starmer and Sugden 2003; Tufano 2010). Although (8) $\bar{b}^p - \bar{b}$ only posts a small deviation, this might suggest that the mean market price does not perfectly converge to either the mean initial belief or the mean post-transaction belief. The subjects' bid or offer might not always be consistent with their own beliefs, which are affected by the transaction price from time to time. Some are more open to making aggressive offers, while some appear hesitant to take the risk. It is observed that heterogeneous subjects can generate different pricing behaviors.[7] If all subjects place orders in accordance with their beliefs, the market price will be fairly close to the mean belief. In other words, when all traders are honest, the market price will be

[7] Several studies have discussed the behaviors of heterogeneous subjects in terms of culture (Oosterbeek, Sloof and Van De Kuilen, 2004; Ehmke, 2006), social preferences (Leibbrandt, 2012; Oechssler, 2012), gender (Dreber, von Essen and Ranehill, 2011; Croson and Gneezy, 2009) and personality (Visser and Roelofs, 2011; Ballinger et al., 2011).

representative of the joint true beliefs.

Nevertheless, we have observed that the market price did not converge to the subjects' beliefs or observations. From our experimental design, the degree of honesty of individual subjects might be further related to the results of the personality traits test. In the following subsection, we will introduce a general model while considering the personality traits for further hypothesis testing.

Variable	K = 10 (n = 10)	K = 20 (n = 13)	K = 30 (n = 13)	K = 40 (n = 20)	K = 50 (n = 11)	K = 60 (n = 20)	K = 70 (n = 12)	K = 80 (n = 12)	K = 90 (n = 9)	All (n = 120)
$p_c - FV$	1.1384 (0.2843)	0.1872 (0.8546)	0.5434 (0.5968)	-1.0711 (0.2975)	1.8770 (0.0900)	**-2.2243** **(0.0385)**	-1.0406 (0.3204)	**-6.2562** **(0.0000)**	**-3.8428** **(0.0049)**	**-2.1511** **(0.0335)**
$p_c - \bar{b}$	0.8207 (0.4330)	-0.4975 (0.6278)	-0.7117 (0.4902)	-0.2571 (0.7999)	0.3891 (0.7053)	**-2.4572** **(0.02379)**	-1.8966 (0.0844)	**-3.6861** **(0.0036)**	**-2.0164** **(0.0785)**	**-2.6399** **(0.0094)**
$p_c - \bar{b}^p$	1.0963 (0.3014)	-0.4003 (0.6959)	-1.1684 (0.2653)	-0.1197 (0.9060)	0.4506 (0.6619)	**-2.3483** **(0.0298)**	-1.4090 (0.1865)	**-2.6771** **(0.0215)**	-1.2641 (0.2418)	**-2.4976** **(0.0139)**
$\bar{b} - FV$	1.1008 (0.2995)	1.1912 (0.2566)	1.3374 (0.2059)	-1.0585 (0.3031)	**2.7436** **(0.0207)**	-0.2220 (0.8267)	-0.2578 (0.8013)	**-2.2881** **(0.0429)**	**-4.1588** **(0.0032)**	-0.2144 (0.8306)
$\bar{b}^p - FV$	0.8740 (0.4048)	0.5829 (0.5708)	1.3724 (0.1950)	-1.1903 (0.2486)	**2.2736** **(0.0463)**	-0.9834 (0.3378)	-0.6632 (0.5209)	**-2.6653** **(0.0220)**	**-4.0714** **(0.0036)**	-1.0157 (0.3118)
$\bar{b} - MB$	**-11.5678** **(0.0000)**	**-4.8543** **(0.0004)**	**-5.2998** **(0.0002)**	**-3.0455** **(0.0067)**	1.5479 (0.1527)	0.7965 (0.4356)	1.6502 (0.1271)	2.0959 (0.0600)	**7.7761** **(0.0000)**	-2.2009 (0.0297)
$\bar{b}^p - MB$	**-5.4848** **(0.0004)**	**-3.2864** **(0.0065)**	**-3.1966** **(0.0077)**	**-2.8385** **(0.0105)**	1.3310 (0.2127)	-0.43616 (0.6676)	0.9474 (0.3638)	0.7130 (0.4906)	1.7565 (0.1171)	-3.2007 (0.0018)

4.2 Panel regression model

In the empirical model, the dependent variables used in the regressions may be associated with two market pricing errors: the first one is the deviation from the fundamental value (K), ($\bar{p} - K$) (fundamental-based error); and the second one is the deviation from the mean post-transaction belief ($\bar{b}^{p}$), ($\bar{p} - \bar{b}^{p}$) (belief-based error). The explanatory variables for both of our regression models are the treatment dummy (D^T), public information dummy (D^p), number of traders (m), gender ratio (g), and the limit order ratio (l) of each round.[8] The gender ratio g is defined by the portion of male subjects. For any individual trader, the limit order ratio tracks the number of limit orders over the total orders placed by the trader. The basic pricing error model using panel estimation for the two error measurements is extended as shown in the following regression:

$$Error = \alpha_t + \beta_1 D_{it}^T + \beta_2 K_{it} + \beta_3 D_{it}^T K_{it} + \beta_4 D_{it}^p + \beta_5 m_i + \beta_6 g_i + \beta_7 l_{it} + \varepsilon_{it}, \qquad (2)$$

where the subscript *it* denotes session *i* of round *t*, and α_t is the round-specific factor, which is a fixed value representing the characteristics of a specific round. The degree of market efficiency can be tested through α_t. The *experience level* of traders may be improved through repeated rounds (t). According to the above setting, we will analyze the round (time) fixed effects regression model.[9]

---

[8] Forsythe et al. (1992) made a similar assumption.

[9] If we analyze the session fixed effects regression model, the effect of some variables which do not change with different rounds is unobserved. As such, no comparison is made between the results of the session fixed effects regression model and pooling model. Besides, the random effects regression model does not clearly account for the effect of increasing familiarity in the trading mechanism on market

4.3 Heterogeneity of personality traits

Personality traits may provide valuable information regarding the individual-specific characteristics of subjects. In our experiment, we did not inflict controls on the traits of subjects participating in any experimental market; instead, we tried to identify possible characteristics of each session, with an attempt to understand the relationship between the heterogeneity of personality traits and market performance. The Five-Factor Model has been applied to our study. It characterizes people based on five dimensions measured by the NEO Personality Inventory in Costa and McCrae (1985), namely, neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness.

The Chinese version of the NEO Personality Inventory test has been made available by Dr. Julian Lai from the City University of Hong Kong.[10] The score range for each dimension is between 10 and 50. We calculated the standard deviation of each personality trait among participants in every session so that the difference between sessions can be captured. These heterogeneous measures of personality traits help

efficiency.

[10] The original personality inventory test and evaluation can be requested from http://www6.cityu.edu.hk/stfprofile/julian.lai.htm

distinguish the following opposite characteristics,

- **Neuroticism**: *sensitive/nervous* as opposed to *secure/confident*
- **Extraversion**: *outgoing/energetic* as opposed to *solitary/reserved*
- **Openness**: *inventive/curious* as opposed to *consistent/cautious*
- **Conscientiousness**: *efficient/organized* as opposed to *easy-going/careless*
- **Agreeableness**: *friendly/compassionate* as opposed to *cold/unkind*

Table 2 provides summary statistics of the variables.

Table 2: Descriptive statistics

Variable	Min.	Max.	Mean	S.D.
Fundamental-based error ($\bar{p}-K$)	-40.58	32.71	-2.12	17.16
Belief-based error ($\bar{p}-\bar{b}^{p}$)	-24.80	33.26	-1.52	8.96
Fundamental value (K)	10	90	49.25	23.70
Number of traders (m)	5	19	12.20	4.17
Gender ratio (g)	0.31	0.88	0.48	0.14
Limit order ratio (l)	0.32	0.73	0.51	0.07
Public information dummy (D^{p})	0	1	0.57	0.50
Mean initial belief ($\bar{b}$)	10	98	49.55	20.88
Mean post-transaction belief ($\bar{b}^{p}$)	8.75	98	48.65	20.48
Mean price ($\bar{p}$)	3.97	80.07	47.13	17.34
S.D. of extraversion	1.47	7.37	5.29	1.34
S.D. of agreeableness	2.91	5.58	3.82	0.67
S.D. of conscientiousness	1.36	6.27	4.66	1.20
S.D. of neuroticism	2.94	8.59	6.19	1.43
S.D. of openness	3.74	6.97	5.25	0.93

4.4 Hypotheses

Based on the pattern of the pricing error, we found that systematic bias is prevalent in the experimental prediction market. Section 4.1 pointed out that error can be sourced from the mean belief and fundamental values. Empirical studies have proved that the number of traders, gender, and limit order ratio are deterministic in this regard. In addition to these variables, we analyzed the characteristics of *public information* and the *heterogeneity of personality traits* in both fundamental- and belief-based models. To be specific, we tried to test if the following hypotheses can withstand scrutiny.

Positive public information will encourage market optimism. In other words,

H_1*: Positive public information leads to market overpricing.*

This study is also connected to empirical results that investigated the variables for the *number of traders*, *gender ratio*, and *limit order ratio* as well as *heterogeneity in personality traits*. The following hypotheses were tested:

H_2*: An increase in the number of traders can improve market efficiency.*

H_3*: An increase in the proportion of male participants can improve market efficiency.*

H_4*: An increase in the ratio of limit orders can improve market efficiency.*

H_5*: An increase in the heterogeneity of personality traits can improve market* *efficiency.*

5 Results and discussion

As can be seen in Tables 3 and 4, both error measurements were significantly influenced by fundamental values. The signs and patterns of the coefficients were consistent with the results presented in Table 1. According to the estimation between the *pooling* and the *fixed effects* regression in Table 3 and Table 4, we observed that there was very little difference between these two models, except for the weakly significant treatment dummy (D^T) of the fixed effects model in Table 3. We will thus focus here on reporting the more informative results of the fixed effects regression. By comparing Model 1' with Model 2'; and Model 3' with Model 4', we found that the coefficients of both error measurements were robust and free from the influence of the heterogeneity of personality traits. The same conclusion can be drawn from the error measurements (1) and (4) in Table 1. By comparing (1) with Model 1' and 2' and (4) with Model 3' and 4', most of the coefficients were significant except for the intercept. In the following, we will discuss the testing results of our earlier hypotheses.

Table 3: Regression results of fundamental-based error

	Pooling		Fixed Effects	
	Model 1	**Model 2**	**Model 1'**	**Model 2'**

Intercept	39.48***	30.01*		
	(11.766)	(14.529)		
D^T	-8.19	-8.32	-7.48*	-7.63
	(5.035)	(5.389)	(4.404)	(4.741)
K	-0.76***	-0.77***	-0.77***	-0.78***
	(0.050)	(0.048)	(0.053)	(0.052)
D^T*K*	0.15*	0.15*	0.13*	0.14*
	(0.083)	(0.085)	(0.076)	(0.081)
D^p	15.51***	13.55***	16.14***	14.15***
	(2.536)	(2.221)	(2.496)	(2.120)
m	-0.88***	-1.03***	-0.88***	-1.04***
	(0.331)	(0.203)	(0.333)	(0.207)
g	10.88	6.74	10.64	6.44
	(11.104)	(8.503)	(11.145)	(8.431)
l	-13.60	-4.70	-14.74	-5.49
	(16.949)	(12.674)	(16.352)	(11.757)
S.D. E		2.65**		2.66**
		(1.050)		(1.060)
S.D. A		2.84		2.85
		(1.980)		(2.013)
S.D. C		-6.23***		-6.19***
		(1.452)		(1.481)
S.D. N		2.70**		2.67**
		(1.278)		(1.308)
S.D. O		-0.39		-0.35
		(1.342)		(1.391)
Adj. R^2	0.63	0.67	0.59	0.63

Notes: *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The robust standard errors are in parentheses.

Table 4: Regression results of belief-based error

	Pooling		**Fixed Effects**	
	Model 3	**Model 4**	**Model 3'**	**Model 4'**
Intercept	12.83	17.32		
	(8.601)	(9.881)		
D^T	-0.86	-1.96	-0.33	-1.49
	(2.709)	(2.770)	(2.631)	(2.641)
K	-0.13***	-0.15***	-0.13***	-0.15***
	(0.020)	(0.021)	(0.025)	(0.026)
D^T*K*	-0.01	0.01	-0.02	0.003
	(0.045)	(0.047)	(0.046)	(0.049)
D^p	-5.39***	-6.87***	-4.67***	-6.12***
	(1.967)	(1.924)	(1.873)	(1.810)
m	-0.39*	-0.37**	-0.39*	-0.38**
	(0.220)	(0.170)	(0.221)	(0.164)
g	6.89	6.57	6.60	6.12
	(7.361)	(5.672)	(7.305)	(5.686)
l	-5.86	-3.16	-5.52	-2.45
	(13.566)	(8.633)	(13.317)	(7.357)
S.D. E		1.96***		1.96***
		(0.654)		(0.660)
S.D. A		0.41		0.49
		(1.423)		(1.484)
S.D. C		-5.22***		-5.11***

		(1.083)		(1.108)
S.D. N		1.99**		1.93**
		(0.813)		(0.856)
S.D. O		-0.76		-0.72
		(1.133)		(1.163)
Adj. R^2	0.37	0.47	0.35	0.45

Notes: *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The robust standard errors are in parentheses.

Table 5: The round fixed effects (α_t) and regression results of their trend

	Model 1'	**Model 2'**	**Model 3'**	**Model 4'**
Round 1	41.28	31.65	13.01	17.17
Round 2	42.00	32.05	15.85	19.59
Round 3	44.10	34.21	14.30	18.20
Round 4	39.07	29.64	10.11	14.27
Round 5	39.90	29.84	12.09	15.95
Round 6	40.47	30.37	12.81	16.44
Round 7	35.46	25.83	10.89	14.93
Round 8	39.05	28.93	12.16	16.00
Intercept	43.29***	33.61***	14.43***	18.45***
	(1.576)	(1.429)	(1.299)	(1.170)
Slope	-0.69*	-0.73**	-0.39	-0.42
	(0.312)	(0.283)	(0.257)	(0.232)
Adj. R^2	0.36	0.45	0.16	0.24

Notes: *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The robust standard errors are in parentheses.

[Figure 2 about here]

5.1 Public information

In Tables 3 and 4, we observed that the coefficients of D^p for the *fundamental-based* error and *belief-based* error have a non-trivial influence on the pricing error albeit in the opposite direction. The average fundamental-based error of positive public information was 16.14% higher than that of the negative case in Model 1' and 14.15%

higher in Model 2’. On the contrary, the positive case caused the average belief-based error to be 4.67% lower than the negative case in Model 3’ and 6.12% lower in Model 4’. This evidence might show that subjects take *public information* to form beliefs and that positive public information has a greater influence than negative public information. In other words, public information first affects the subjects’ beliefs and then the market price. By comparing the fundamental-based and belief-based errors, it is found that market price reacts toward the direction of public information but less than the magnitude of the mean belief, which represents the subjects’ original thoughts.

5.2 Number of traders

The effect of the number of traders was significant. Pricing error can be reduced by 0.88% (Model 1’) and 1.04% (Model 2’) per participant in terms of the fundamental-based error and 0.39% (Model 3’) and 0.38% (Model 4’) in terms of the belief-based error. Since participants carry independently observed information, the increase in the number of traders can be beneficial to the market’s aggregate information. Please note that this result is based on our sampling observation; the actual relationship may, however, not take a linear form between the number of traders and price efficiency (Gresik and Satterthwaite 1989).

5.3 Gender difference

As can be seen in Table 2, the gender ratio does not remain constant, with the actual male ratio ranging from 31% to 88%. Tables 3 and 4 show that the effect of the gender ratio on price efficiency has disappeared. Unlike political contracts of the IEM, where male participants demonstrate more aspiration in participation and outward preference expression, the effect of the gender ratio in terms of the individual preference for a specific contract is ignored in our experimental market. This might be because subjects are motivated by monetary reward; their goal is only to seek profitable opportunities as much as possible.

5.4 Limit order ratio

Table 2 shows that the mean limit order ratio per round was 0.51, with a standard deviation of 0.07. Since the trading period was restricted, bidding strategies were also conditional on the time constraint. Most participants applied both market and limit orders as much as they could in order to seize profitable opportunities. As such it is not surprising that in Tables 3 and 4 the information regarding the limit order ratio did not improve market efficiency.

5.5 Heterogeneity of personality traits

According to the regression results of Model 2’ and Model 4’ in Tables 3 and 4, the influence of personality traits was concentrated on three aspects, which were extraversion, conscientiousness, and neuroticism. The coefficients of these variables were significant and consistent in the fundamental- and belief-based models. Among the three traits, the heterogeneity of conscientiousness was the only one that improved pricing efficiency. This result implies that the percentage of *efficient* and *organized* traders in the market is important. The heterogeneity of extraversion and neuroticism might undermine market efficiency possibly because *sensitive*, *nervous*, *energetic* or *solitary* individuals might find it challenging to make or accept offers based on their information or belief.

5.6 The degree of experience with repeated trading

From Figure 2 and Table 5, we observe that the trend of the fixed effects is declining, being higher in the first few rounds, which suggests that subjects can become familiar with the trading mechanism. As a result, this may improve the efficiency of the experimental prediction market. This phenomenon is generally called “the experience effect”. From Table 3 to Table 5, the results of the pooling model are coherent with those of the fixed effects model, implying that our analysis is robust and exonerated from the experience effect. The fixed effects model can be regarded as a robustness check for the pooling model after considering the experience effect. Figure 2 shows the belief-based error model’s trend for each round. Table 5 shows that the slopes are significant in Models 1’ and 2’ under fundamental-based fittings.

6 Conclusion

This study discusses whether the prediction market can achieve efficiency when participants have heterogeneous beliefs and personality traits. Through a series of experiments under the framework of a double auction futures market, it is found that as a medium to communicate and aggregate individual information, the market is hardly efficient even if manipulators are removed. A number of studies have shown that efficiency might be compromised by a number of factors (see Ottaviani and Sørensen, 2007). In this paper, we try to measure the influence of possible factors that cause market failure in the hope of increasing efficiency in the prediction market.

Based on our analysis of the experimental results, there are five possible causes of market inefficiency. First, price is largely influenced by *public information*. Regardless of initial beliefs, participants tend to move in the same direction as what has been disclosed publicly. Second, the larger the number of participants, the more accurate the prediction market can be. Each individual owns a certain amount of information. Through transactions, such information is naturally shared with others, regardless of whether he/she likes it or not. Third, the gender ratio might not be able to improve market efficiency in an experimental market mainly driven by monetary reward. Fourth, no significant bidding behavior can affect market efficiency. Last but not least, personality is found to be a deterministic factor of market efficiency. Participants may hold different views even though they are exposed to identical information. To summarize, three out of the above five personality traits might have some bottom up influence on prices and henceforth the efficiency of the prediction market.

This study demonstrates how individual beliefs can shape or bias the market through aggregation. The relationship between heterogeneous personality and market joint behaviors has also been revealed. This discussion is unprecedented in the experimental prediction market and deserves to receive more attention in future studies.

References

Abdellaoui, M. (2000). Parameter-free elicitation of utilities and probability weighting functions. *Management Science*, 46, 1497–1512.

Abdellaoui, M., F. Vossmann, & M. Weber (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Science*, 51, 1384–1399.

Ackert, L. F., B. K. Church, & P. Zhang (2002). Market behavior in the presence of divergent and imperfect private information: experimental evidence from Canada, China, and the United States. *Journal of Economic Behavior & Organization*, 47, 435–450.

Ali, M. M. (1977). Probability and utility estimates for racetrack bettors. *Journal of Political Economy*, 85, 803–15.

Ballinger, T. P., E. Hudson, L. Karkoviata, & N. T. Wilcox (2011). Saving behavior and cognitive abilities. *Experimental Economics*, 14, 349–374.

Berg, J., F. Nelson, & T. Rietz (2008). Prediction market accuracy in the long run. *International Journal of Forecasting*, 24, 285–300.

Berg, J. & T. Rietz (2006). The Iowa electronic markets: stylized facts and open issues. in R. W. Hahn and P. C. Tetlock (ed.), *Information Markets: A New Way of Making Decisions*, AEI-Brookings Joint Center for Regulatory Studies, 142-169.

Camerer, C. F. & T. H. Ho (1994). Nonlinear weighting of probabilities and violations of the betweenness axiom. *Journal of Risk and Uncertainty*, 8, 167-196.

Chamberlin, E. H. (1948). An experimental imperfect market. *Journal of Political Economy*, 56, 95–108.

Christiansen, J. D. (2007). Prediction markets: practical experiments in small markets and behaviours observed. *The Journal of Prediction Markets*, 1, 17–41.

Copeland, T. E. & D. Friedman (1987). The effect of sequential information arrival on asset prices: an experimental study. *Journal of Finance*, 42, 763–797.

Costa, P. T. & R. R. McCrae Jr. (1985). *The NEO personality inventory manual*, Odessa, FL: Psychological Assessment Resources.

Croson, R. & U. Gneezy (2009). Gender differences in preferences. *Journal of Economic Literature*, 47, 448–474.

Deck, C., S. Lin, & D. Porter (2013). Affecting policy by manipulating prediction markets: experimental evidence. *Journal of Economic Behavior & Organization*, 85, 48–62.

Dreber, A., E. von Essen, & E. Ranehill (2011). Outrunning the gender gap—boys and girls compete equally. *Experimental Economics*, 14, 567–582.

Ehmke, M. D. T. (2006). Dissertation abstract: The influence of culture on economic behavior with applications to food and the environment. *Experimental Economics*, 9, 167–168.

Figlewski, S. (1978). Market ‘efficiency’ in a market with heterogeneous information. *Journal of Political Economy*, 86, 581–597.

Figlewski, S. (1979). Subjective information and market efficiency in a betting market. *The Journal of Political Economy*, 87, 75–88.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10, 171–178.

Forsythe, R. & R. Lundholm (1990). Information aggregation in an experimental market. *Econometrica*, 58, 309–347.

Forsythe, R., F. Nelson, G. Neumann, & J. Wright (1992). Anatomy of an experimental political stock market. *The American Economic Review*, 82, 1142–1161.

Forsythe, R., T. Rietz, and T. Ross. (1999). Wishes, expectations and actions: a survey on price formation in election stock markets. *Journal of Economic Behavior & Organization*, 39, 83–110.

Gjerstad, S. (2005). Risk aversion, beliefs, and prediction market equilibrium. *Unpublished Manuscript, Economic Science Laboratory*, University of Arizona.

Gresik, T. A. & M. A. Satterthwaite (1989). The rate at which a simple market converges to efficiency as the number of traders increases: an asymptotic result for optimal trading mechanisms. *Journal of Economic Theory*, 48, 304–332.

Griffith, R. M. (1949). Odds adjustments by american horse-race bettors. *The American Journal of Psychology*, 62, 290–294.

Hahn, R. W. & P. C. Tetlock (2006). *Information markets: a new way of making decisions*, AEI-Brookings Joint Center for Regulatory Studies.

Hanson, R. D. (2006). Designing real terrorism futures. *Public Choice*, 128, 257–274.

Hanson, R., R. Oprea, & D. Porter (2006). Information aggregation and manipulation in an experimental market. *Journal of Economic Behavior & Organization*, 60, 449–459.

Haruvy, E., Y. Lahav, & C. N. Noussair (2007). Traders' expectations in asset markets: experimental evidence. *American Economic Review*, 97, 1901–1920.

Hayek, F. (1945). The use of knowledge in society. *American Economic Review*, 35, 519–530.

Hazen, T. L. (1987). Volatility and market inefficiency: a commentary on the effects of options, futures, and risk arbitrage on the stock market. *Washington and Lee Law Review*, 44, 789–805.

Hurley, W. & L. McDonough (1995). Note on the Hayek hypothesis and the favorite-longshot bias in parimutuel betting. *American Economic Review*, 85, 949–955.

Isaacs, R. (1953). Optimal horse race bets. *American Mathematical Monthly*, 60, 310–315.

Knetsch, J., F. F. Tang, & R. Thaler (2001). The endowment effect and repeated market trials: is the Vickrey auction demand revealing? *Experimental Economics*, 4, 257–269.

Lee, D. S. & E. Moretti (2009). Bayesian learning and the pricing of new information: evidence from prediction markets. *American Economic Review*, 99, 330–336.

Leibbrandt, A. (2012). Are social preferences related to market performance? *Experimental Economics*, 15, 589–603.

Loomes, G., C. Starmer, & R. Sugden (2003). Do anomalies disappear in repeated markets? *Economic Journal*, 113, C153–C166.

Maloney, M. T. & J. H. Mulherin (2003). The complexity of price discovery in an efficient market: the stock market reaction to the Challenger crash. *Journal of Corporate Finance*, 9, 453–479.

Oechssler, J. (2012). Finitely repeated games with social preferences. *Experimental Economics*, doi:10.1007/s10683-012-9336-6.

Oliven, K. & T. A. Rietz (2004). Suckers are born but markets are made: individual rationality, arbitrage, and market efficiency on an electronic futures market. *Management Science*, 50, 336–351.

Oosterbeek, H., R. Sloof, & G. Van De Kuilen (2004). Cultural differences in ultimatum game experiments: evidence from a meta-analysis. *Experimental Economics*, 7, 171–188.

Ottaviani, M. & P. N. Sørensen (2007). Outcome manipulation in corporate prediction markets. *Journal of the European Economic Association*, 5, 554–563.

Ottaviani, M. & P. N. Sørensen (2010). Noise, information, and the favorite-longshot bias in parimutuel predictions. *American Economic Journal: Microeconomics*, 2, 58–85.

Plott, C. R. & S. Sunder (1982). Efficiency of experimental security markets with insider information: an application of rational-expectations models. *Journal of Political Economy,* 90, 663–698.

Plott, C. R. & S. Sunder (1988). Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica*, 56, 1085–1118.

Plott, C. R., J. Wit, & W.C. Yang (2003). Parimutuel betting markets as information aggregation devices: experimental results. *Economic Theory*, 22, 311–351.

Roll, R. (1984). Orange juice and weather. *American Economic Review*, 74, 861–880.

Rosenberg, B., K. Reid, & R. Lanstein (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management*, 11, 9–16.

Smith, V. L. (1962). An experimental study of competitive market behavior. *Journal of Political Economy*, 70, 111–137.

Smith, V. L. (1982). Microeconomic systems as an experimental science. *American Economic Review*, 72, 923–955.

Snowberg, E., J. Wolfers, & E. Zitzewitz (2007). Partisan impacts on the economy: evidence from prediction markets and close elections. *Quarterly Journal of Economics*, 122, 807–829.

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk* *and Uncertainty*, 32, 101–130.

Terrell, D. & A. Farmer (1996). Optimal betting and efficiency in parimutuel betting markets with information costs. *Economic Journal*, 106, 846–868.

Tufano, F. (2010). Are ‘true’ preferences revealed in repeated markets? an experimental demonstration of context-dependent valuations. *Experimental Economics*, 13, 1–13.

Tversky, A. & C. Fox (1995). Weighing risk and uncertainty. *Psychological Review*, 102, 269–283.

Tversky, A. & D. Kahneman (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.

Veiga, H. & M. Vorsatz (2010). Information aggregation in experimental asset markets in the presence of a manipulator. *Experimental Economics*, 13, 379–398.

Visser, M. S. & M. R. Roelofs (2011). Heterogeneous preferences for altruism: gender and personality, social status, giving and taking. *Experimental Economics*, 14, 490–506.

Weitzman, M. (1965). Utility analysis and group behavior: an empirical study. *Journal of Political Economy*, 73, 18–26.

Wlezien, C. & R. S. Erikson (2004). The fundamentals, the polls, and the presidential vote. *PS: Political Science and Politics*, 37, 747–751.

Wolfers, J. & E. Zitzewitz (2004). Prediction markets. *The Journal of Economic Perspectives*, 18, 107–126.

Wu, G. & R. Gonzalez (1996). Curvature of the probability weighting function. *Management Science*, 42, 1676–1690.

Appendix A: Experimental design

Table A1: Experimental design

treatment	**market**	**number of traders (*m*)**	**fundamental value (*K*)**	**public information (*y*)**	**private information (*x*)**	**settlement price (*S*)**
WTA_0	double auction	[5,19]	$K < 50$	[0,1]	[0,2]	0
WTA_50	double auction	[5,19]	$K = 50$	[0,1]	[0,2]	50
WTA_100	double auction	[5,19]	$K > 50$	[0,1]	[0,2]	100
Ratio	double auction	[5,19]	[10,90]	[0,1]	[0,2]	K

Table A2: Session description

session	treatment sequence	number of traders (m)	fundamental value (K)	public information (y)	average observation ($x + y$)	settlement price (S)
1	Ratio	16	20, 40, 60, 30	0, 0, 1, 0	10, 52, 77, 23	20, 40, 60, 30
	WTA		30, 60, 40, 20	0, 0, 0, 0	21, 35, 21, 23	0, 100, 0, 0
2	WTA	15	60, 40, 80, 70	1, 1, 1, 1	84, 53, 82, 76	100, 0, 100, 100
	Ratio		70, 40, 60, 80	1, 1, 1, 1	44, 56, 78, 87	70, 40, 60, 80
3	Ratio	12	40, 80, 70, 30	1, 1, 1, 1	64, 78, 89, 83	40, 80, 70, 30
	WTA		30, 40, 60, 20	0, 1, 1, 0	8, 58, 64, 17	0, 0, 100, 0
4	WTA	19	60, 80, 70, 40	1, 1, 1, 1	81, 93, 86, 63	100, 100, 100, 0
	Ratio		20, 30, 40, 60	0, 1, 0, 1	9, 60, 35, 70	20, 30, 40, 60
5	Ratio	15	70, 60, 40, 80	0, 1, 0, 1	42, 71, 24, 87	70, 60, 40, 80
	WTA		40, 30, 20, 60	1, 1, 0, 1	53, 42, 24, 73	0, 0, 0, 100
6	WTA	8	20, 60, 40, 30	1, 0, 0, 1	42, 29, 42, 54	0, 100, 0, 0
	Ratio		30, 20, 40, 60	0, 1, 1, 0	29, 63, 50, 29	30, 20, 40, 60
7	Ratio	6	80, 40, 60, 70	1, 0, 0,1	78, 39, 50, 94	80, 40, 60, 70
	WTA		40, 20, 30, 60	0, 0, 1, 1	28, 17, 50, 78	0, 0, 0, 100
8	WTA	5	60, 40, 70, 80	1, 0, 1, 1	73, 27, 87, 87	100, 0, 100, 100
	Ratio		60, 20, 30, 40	1, 0, 0, 0	67, 13, 27, 27	60, 20, 30, 40
9	Ratio	13	40, 20, 30, 60	0, 0, 1, 1	36, 13, 49, 74	40, 20, 30, 60
	WTA		60, 70, 80, 40	1, 1, 1, 0	79, 79, 82, 33	100, 100, 100, 0
10	Ratio	18	80, 70, 60, 40	0, 1, 1, 0	54, 83, 72, 37	80, 70, 60, 40
	WTA		30, 60, 20, 40	0, 1, 0, 0	24, 76, 9, 30	0, 100, 0, 0
11	WTA	8	10, 50, 10, 90	0, 0, 0, 1	4, 42, 8, 88	0, 50, 0, 100
	Ratio		50, 10, 90, 20	1, 0, 1, 0	71, 8, 83, 21	50, 10, 90, 20
12	Ratio	11	90, 10, 30, 50	1, 0, 0, 1	85, 27, 24, 67	90, 10, 30, 50
	WTA		90, 50, 70, 50	1, 1, 1, 1	91, 70, 82, 67	100, 50, 100, 50

Table A2 (*Continued*)

session	treatment sequence	number of traders (m)	fundamental value (K)	public information (y)	average observation ($x + y$)	settlement price (S)
13	WTA	11	80, 50, 10, 90	1, 1, 0, 1	88, 67, 3, 91	100, 50, 0, 100
	Ratio		90, 10, 50, 90	1, 0, 1, 1	94, 6, 61, 94	90, 10, 50, 90
14	Ratio	10	90, 50, 80, 10	1, 0, 0, 0	97, 30, 40, 7	90, 50, 80, 10
	WTA		10, 50, 20, 70	0, 1, 0, 1	7, 77, 17, 77	0, 50, 0, 100
15	WTA	16	10, 90, 80, 10	0, 1, 1, 0	4, 94, 79, 13	0, 100, 100, 0
	Ratio		50, 30, 50, 70	1, 0, 1, 0	63, 15, 52, 46	50, 30, 50, 70

Appendix B: Subject Instruction

1 Instruction

This is an experiment of predicting a coming event. You will be paid in cash for your participation at the end of the experiment. Different participants may earn different reward. What you earn depends on your decision and that of the others.

The experiment will take place at the computer terminal you are seated. If you have any question during the instruction period, please raise your hand and an assistant will stop by to answer your question. If any difficulties arise after the experiment has begun, raise your hand, and someone will assist you.

In this experiment your role is a **Trader** in the market. There will be **eight market rounds**. Each market round is separate and will last for **3 minutes**. Your experimental earning will be the cumulative sum of what you earned in each market round. At the end of the day, your experimental earning will be normalized so that your final payoff will fall into the range between **NT$260** and **NT$500**.

2 Information

2.1 Your Information

Here is a black box that has **10 balls** in it, ***k*** yellow balls and **10 – *k*** white balls. Before each round begins, all traders will observe their own private information through drawing two balls one at a time with replacement from the black box. Each trader makes observation privately so that no one will know what information anyone else has observed. After all traders finish their turns of drawing, the experimenter will also draw a ball with replacement from the black box and make public its color. Attention, all 10 balls remain unchanged in the black box. **Your mission is to predict the number of yellow balls (*k*)**. Based on your observation, your information could be zero, one, two or three yellow balls.

After all participants finishing drawing and observation, each trader has to answer the two questions below, which will be shown on your screen [see Figure B1].

1. **How many yellow balls did you observe, including your draw and public announcement?**

2. **What do you believe is the percentage of yellow balls contained in the black box?**

A trading period is launched immediately after all traders complete these two questions.

Figure B1: Your observation and belief

2.2 Double Auction

Each round, traders will be lending **1,000** tokens from the experimenter. Traders are allowed to buy or sell with these tokens. At the end of the market round, the 1,000 tokens will be returned to the experimenter. **Your goal is try to use all available information to buy low and sell high. The more tokens you make the more payoff you will earn in New Taiwan Dollars.** Double auction is adopted as the trading mechanism in this experimental market.

2.2.1 The Interface

Now we will describe a trader's screen [see Figure B2].

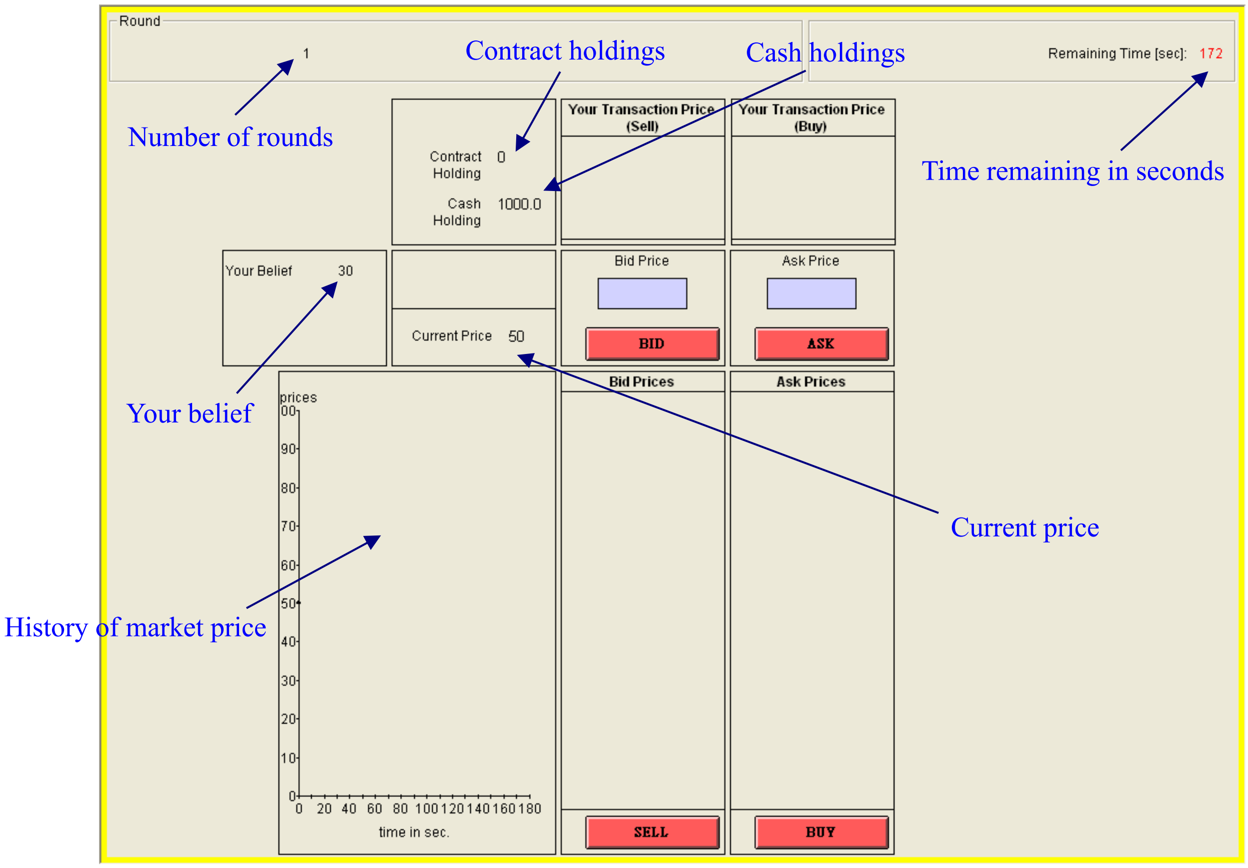


Figure B2: Trader's screen

Market price represents the market opinion of the percentage of yellow ball. The price starting at 50 means there are "five" yellow balls in the black box, the maximum uncertainty on the percentage of yellow ball. You can think of this like flipping a coin; if it is head the event is yellow and if it lands on tail the event is white.

Your contract holding position can be positive or negative. Positive means net buying, otherwise means net selling. Your cash holding is just temporary, for that all contract holding will also be settled with cash at the end of each round.

2.2.2 Trading Illustration

If you think the proportion of yellow ball is higher than the current price, you can submit a bid order and press the **bid** button. Your bid order will then be shown under the bid column. In like manner, if you think the proportion of yellow ball is lower than the current price, you can submit an ask order and press the **ask** button. Then your ask order will be shown under the ask column accordingly. The orders you placed will be denoted in blue and the others will be in black.

Sending your own bid or ask order is called "limit order". You can also directly accept orders of others, which is called "market order". If you wish to buy (sell), you can press the **BUY** (**SELL**) button to accept the lowest ask order (highest bid order) in the market. Each transaction price will be shown on your screen above bid and ask column.

2.3 After Trading

Your belief may change during the process of transaction. Please key in your Post-Transaction belief of the proportion of yellow ball in the black box [see Figure B3].

The closer of your Post-Transaction belief the more tokens you will earn. The reward policy is

50 – │Your Post-Transaction Belief – Real Proportion of the Yellow Ball│

Figure B3: Post-Transaction belief and reward policy

2.4 Settlement

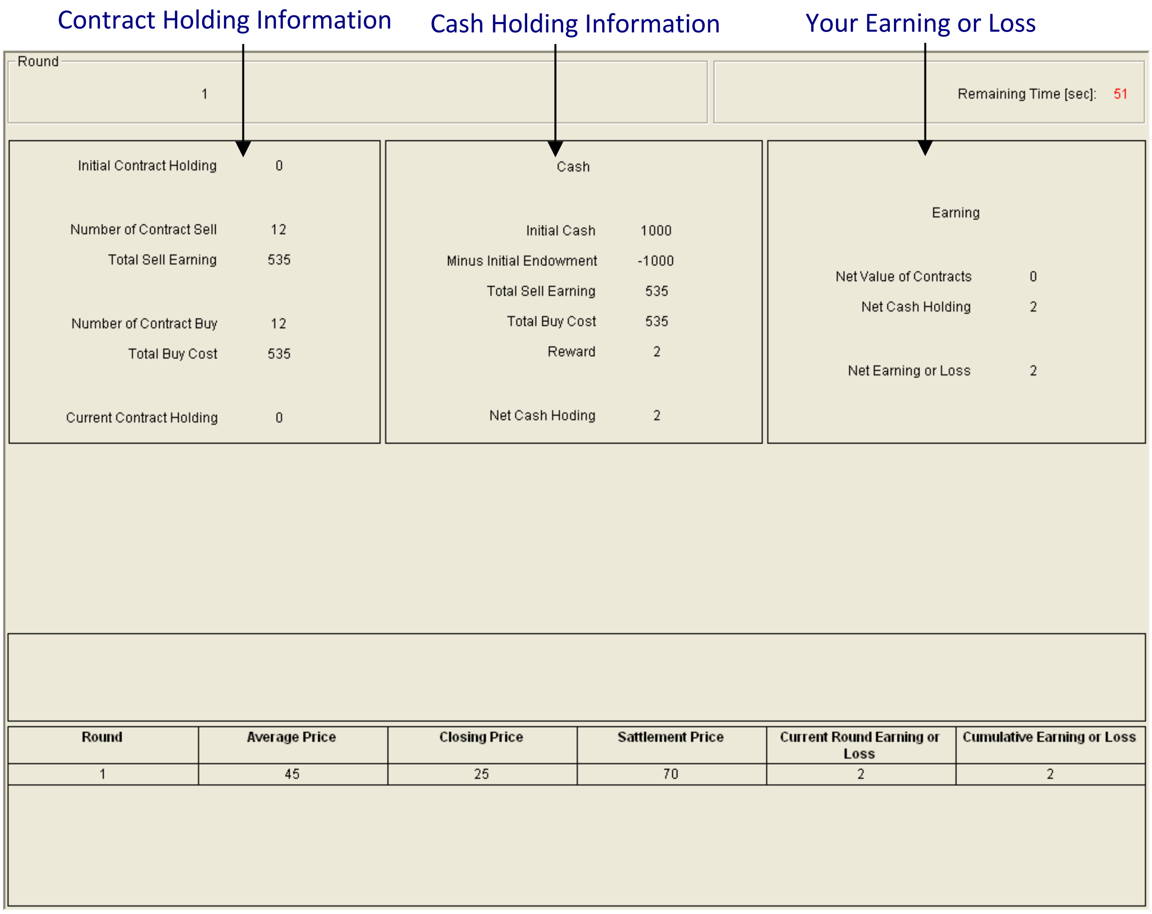


Figure B4: Settlement information

The market adopts cash settlement. For settlement price (***S***) and your contract holding position (***h***), you will receive (***S*** x ***h***) in token value. Here, ***S*** is dependent on the type of two treatments applied. For winner-take-all treatment, the settlement price ***S*** is determined as below:

$$S = 100, \text{ if } k > 5;\ S = 50, \text{ if } k = 5; \text{ and } S = 0, \text{ if } k < 5.$$

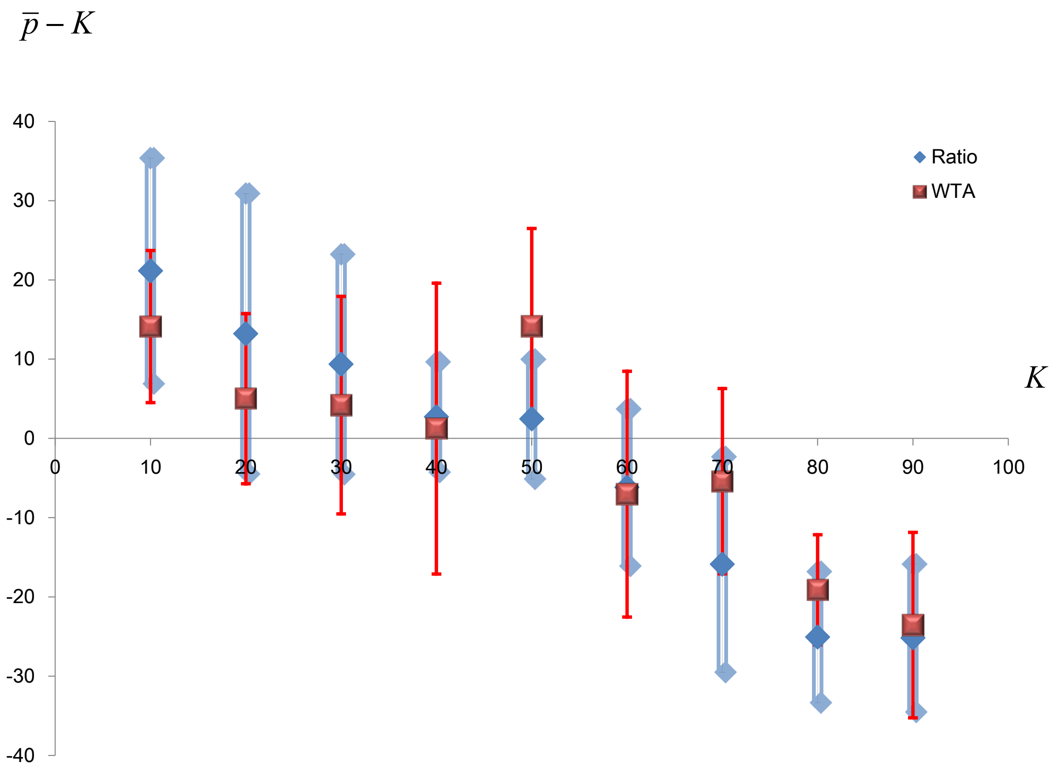
For ratio treatment, the settlement price ***S*** is defined as $S = 10 \times k$.

The settlement process takes place after all traders declare their Post-Transaction beliefs.

3. Personality Test

Please complete the personality test based on your current thoughts at comfortable pace. Your answer will be treated as confidential information and will not be used for any other purpose. Your answer will improve the understanding of the relationship between decision-making and personality.

Figure 1: Fundamental-based pricing error versus fundamental value



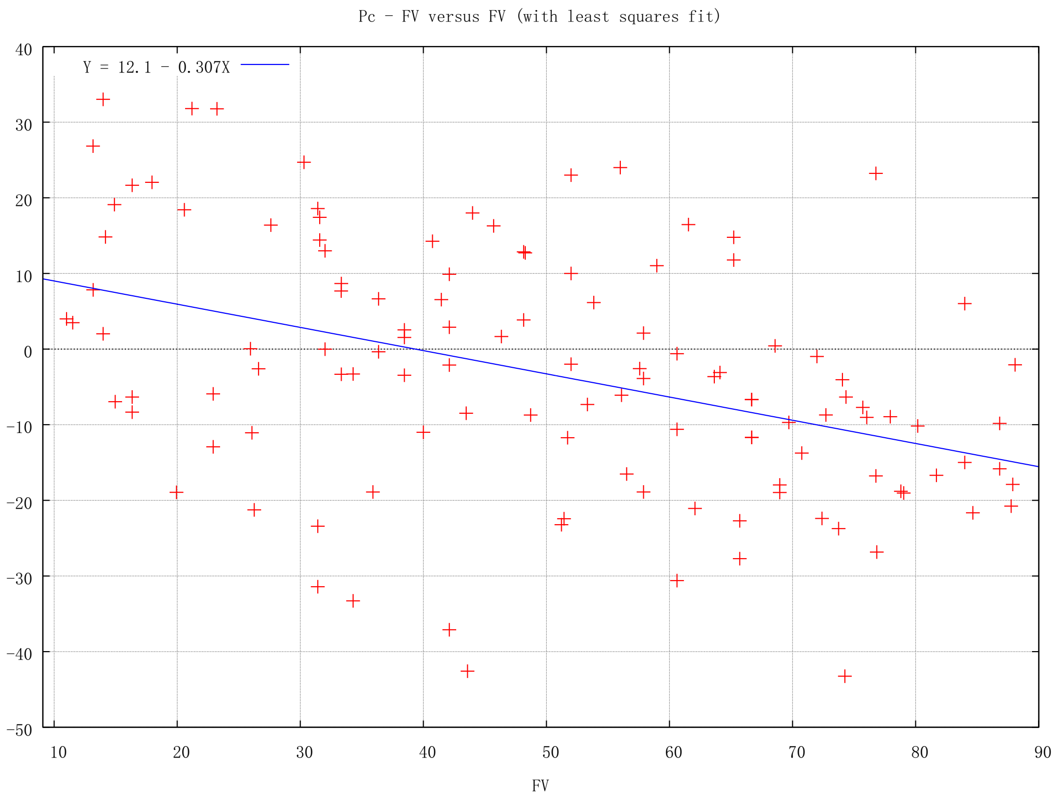

Pc - FV versus FV (with least squares fit)
Y = 12.1 - 0.307X
40
30
20
10
0
-10
-20
-30
-40
-50
10
20
30
40
50
60
70
80
90
FV

Figure 2: The trend of round fixed effects (α_t)