

Reconsidering Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets

Brice Corgnet	Cary Deck	Mark DeSantis	Kyle Hampton	Erik O. Kimbrough
Emlyon Business School	University of Alabama & Chapman University	Chapman University	Chapman University	Chapman University

January 2020

Abstract

The ability of markets to aggregate dispersed information is a cornerstone of economics and finance. In a seminal experiment, Plott and Sunder (1988) offer support for the rational expectations hypothesis. However, recent laboratory experiments have called into question the robustness of those initial results. In this paper, we offer the first attempt to directly replicate key findings of the original study. Failing to replicate their results, the post-study probability that market performance is better described by rational expectations than the prior information (Walrasian) model is low. Given this result, we introduce a new treatment that implements a market structure consistent with naturally occurring prediction markets, which can be viewed as completing the original experimental design. In this new treatment, we find strong support for the rational expectations model. Thus, while the original paper showed conditions where markets *can* aggregate information, we attempt to identify *sufficient conditions* for such aggregation to be robust.

Keywords: Information Aggregation, Efficient Markets, Rational Expectations, Replication

Acknowledgements: Brice Corgnet performed this research within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d’Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR). The research was conducted with the approval of Chapman University’s Institutional Review Board, application #IRB-19-29.

1. Introduction

The ability of markets to aggregate dispersed information has long played a critical role in the debate over the optimal allocation of scarce resources in Economics. Hayek (1945) famously emphasized that markets could serve as a mechanism to aggregate and convey highly-dispersed information to the rest of society via prices. Although most economists are sympathetic to the Hayekian argument, the ongoing debate between classical and behavioral finance testifies to the existence of open disagreements regarding the extent to which the *efficient market hypothesis* holds in practice, i.e., the extent to which market prices accurately reflect available information (Fama, 1970; Fama, 2008; Shleifer, 2000; Thaler, 2005, 2015; Shiller, 2015). This debate is fueled by the lack of a widely-accepted methodology to gauge whether naturally occurring markets are indeed efficient. Indeed, Fama (1991) acknowledged that “market efficiency per se is not testable” (p. 1575).

However, the emergence of experimental economics (Chamberlain, 1948; Smith, 1962) made it possible for economists to construct markets in which they would not only have perfect control over the distribution of private information but also know the fundamental value of the asset being traded. Thus, experimental markets enable researchers to study the aggregation of private information, rendering the *efficient market hypothesis* testable. This effort was initiated by Plott and Sunder (1988), hereafter P&S, who designed experimental markets to gauge the extent to which dispersed private information was incorporated into prices.¹ In their design, subjects traded assets that delivered one of three possible dividends at the end of each market period based on a randomly determined state (X, Y, or Z). Each trader was endowed with private information ruling out one of the possible states (e.g., Not X). Because P&S ensured half the traders were given one piece of information, while the other half were given the complementary piece of information (e.g., half told not X and half told Not Y), the aggregate information held by market participants was complete (e.g., the state must be Z).

P&S report evidence that, in line with the *efficient market hypothesis*, market prices *can* reflect the true asset value if there are contingent markets for each possible state (referred to in P&S as Series B) or if there is a single asset yielding common state-specific dividends to all traders (referred to in P&S as Series C). An equally important, but oft-ignored finding of P&S, is the fragility of successful information aggregation. P&S (p. 1117) highlight this in their conclusion when they point out that “not all markets can be depended upon to behave in accord with the rational expectations model.” In particular, P&S found that markets with a single asset but diverse state-specific dividends (referred to in P&S as Series A) do not yield prices consistent

¹ In a previous paper, Plott and Sunder (1982) demonstrate that markets disseminate information as prices reflect the information of fully informed insiders.

with the rational expectations model.² Even for Series B and Series C, P&S are careful to qualify their conclusions as “demonstrat[ing] that markets *can* aggregate diverse information in a manner consistent with rational expectations models” (P&S, p. 1116, emphasis added) and not that they will. However, the literature has often taken P&S to imply that markets *will* aggregate information. For example, Choo, Kaplan, and Zultan (2019) state “[P&S] show that market prices converge to the true value of the asset” (p. 627). Kwasnica, Velthuis, and Williams (2019) write that P&S “document that the rational expectations equilibrium provides an accurate prediction of prices” (p. 329). Rud, Rabanal, and Sharifova (2019) cite P&S when noting “competitive markets aggregate information quickly” (p. 263).³

The positive results of P&S have been highly influential and prediction markets have subsequently come to be viewed as a mainstream tool for decision makers wanting to aggregate information (see Arrow et al., 2008). There has been a wave of field studies discussing the successful use of prediction markets at major companies such as Hewlett-Packard, Ford, Google, Microsoft, Yahoo and IBM (Chen and Plott, 2002; Cowgill, Wolfers and Zitzewitz, 2009; Cowgill and Zitzewitz, 2015).⁴ Pennock et al., (2001) and Gürkaynak and Wolfers (2006) find that prediction markets are better at forecasting Oscar winners and economic variables, respectively, than experts in those areas. Tetlock (2004) demonstrates that financial and political contracts on (the now shuttered) Tradesports.com were efficient.⁵ Berg, Forsythe, Nelson, and Rietz (2008) report that the Iowa Electronic Markets outperform more traditional polling. More recently, Atanasov, et al., (2017) show that prediction markets were more accurate than the means of prediction polls for a wide variety of 261 economic and policy event forecasts.

Despite these affirmative empirical results, evidence from a number of recent experimental studies has cast doubt on the robustness of the market’s ability to aggregate dispersed private information. Several works using variants of the P&S design with a single asset (Series C) have reported substantial mispricing (see Biais, Hilton, Mazurier and Pouget, 2005; Hanson, Oprea and Porter 2006; Veiga and Vorsatz, 2010; Huber, Angerer and Kirchler, 2011; Corgnet, DeSantis and Porter, 2015, 2019; Page and Siemroth, 2017; Corgnet, Deck, DeSantis and Porter, 2018). Similarly, Page and Siemroth (2018) report a failure to aggregate private information in contingent claims markets, though their setup differed from P&S (Series B), as they studied the

² A follow-up study by Forsythe and Lundholm (1990) showed that full information aggregation could be achieved when the same cohort of subjects was re-invited for a second trading session where the distribution of heterogeneous dividend values of the single asset was common knowledge among all traders.

³ There are many papers that use somewhat ambiguous phrasing on this point, but certainly do not warn readers that P&S shows aggregation is possible rather than robust.

⁴ There have also been several subsequent studies examining the ability of other systems to aggregate diverse information (Hampton 2007; Gillen, Plott and Shum, 2017; Court, Gillen, McKenzie and Plott, 2018).

⁵ Sports contracts on the exchange exhibited some inefficiencies such as a bias for longshots as has been observed in other gambling data.

more intricate case of aggregate uncertainty (see Anderson and Holt, 1997) in which all private signals available in the market cannot fully reveal the true asset value.⁶ Of course, there have been some studies supporting the rational expectations hypothesis. For example, Choo, Kaplan, and Zultan (2019) observed information aggregation in a setting with common state-specific dividends and a full set of contingent claims. In this setting, the information structure differs from P&S as traders possessing different private signals may hold identical beliefs about the true asset value.

This paper seeks to reconcile the apparent contradictory evidence regarding the ability of markets to aggregate information. Mixed evidence in a given field is likely to produce self-serving biases leading researchers to interpret the ambiguous evidence as supporting their own views of the world (Babcock and Loewenstein, 1997; Dawson, Gilovich and Reagan, 2002; Maniadis, Tufano and List, 2014). Progress thus requires the establishment of empirical facts that outline critical boundary conditions of the informational efficiency of markets. To that end, we conduct a direct replication of P&S Series B (contingent claims) and C markets (homogeneous values).⁷ Replication has been shown to be especially relevant to cases in which original studies report striking findings based on small samples as in P&S (De Long and Lang, 1992; Ioannidis, 2005; Maniadis, Tufano and List, 2014).⁸ This is the case because the probability that a declared research finding is actually true (the so-called 'Post-Study Probability' or PSP in Maniadis, Tufano and List, 2014) increases with sample size and the prior belief in the hypothesis. Although prior beliefs (before P&S reported their findings) regarding the empirical success of rational-expectation models are difficult to pin down, the precursory statement of Radner (1982, p. 942) that these models require computational capacity "*far beyond what is realistic*" when traders hold dispersed information suggest they were low. However, given the magnitude of the effects observed in P&S, all but the most ardent skeptics would be convinced by the evidence put forth in P&S.

Ultimately, we fail to replicate the results of P&S. Given the statistical power of our replication, the PSP that Series B- and C-style markets aggregate information is very low even if one held strong affirmative priors before observing the seemingly compelling evidence in P&S. This result highlights the importance of interpreting P&S as providing a sort of existence proof for efficient markets rather than robust evidence that markets in general are efficient. It also emphasizes the need for independent replication of results.

⁶ Page and Siemroth (2018) consider a common dividend value environment.

⁷ Previous partial replications have introduced potential confounds such as computerization (Hanson, Oprea and Porter, 2006), the use of call auctions (Biais, Hilton, Mazurier and Pouget, 2005), and a cost of information acquisition (Corgnet, Deck, DeSantis and Porter, 2018).

⁸ The main findings of P&S regarding information aggregation in series B and C are based on only five markets each of which uses a different information structure.

Given our finding, we then extended the experimental design of P&S by introducing Series D markets in which traders have homogeneous values *and* there are contingent claims markets.⁹ Many of the successful prediction markets in operation discussed above as well as the laboratory experiments of Choo, Kaplan and Zultan (2019) are best described as Series D markets; however, we are not aware of any studies directly comparing Series D markets to Series B or C markets.¹⁰ In our Series D markets we find overwhelming support for the rational expectations hypothesis and the PSP is quite high even for those who are skeptical a priori.

Taken together, our results and those of P&S indicate that whether markets aggregate information or not is to a large degree a matter of market design. Markets can be structured in ways that are more or less conducive to information aggregation. In particular, the evidence suggests that homogeneous values and contingent claims, features that are present in most prediction markets in the field, both enhance the likelihood of information aggregation. Further, it is the combination of these two features that yields reliable information aggregation.

2. The Experiments of Plott and Sunder

In P&S traders are endowed with an amount of money, denoted in Francs, and assets called certificates that can be bought and sold by the traders using a continuous double auction. At the end of each trading period, certificates pay their owner a dividend that depends on the state of nature, which can be X, Y, or Z. The state of nature is not known by any trader, but the probability distribution over states is common information. Further, half the traders are informed of one unrealized state; while, the other half are informed of the other unrealized state. So for example, if the state is X, then half the traders are informed the state is not Y and half are informed it is not Z. In aggregate the traders have sufficient information to determine that the state is X. If markets successfully aggregate dispersed information, then the price should reflect the value the traders have for the certificates in the realized state.

P&S present the results of 11 total sessions from three series of markets. Although never described as such, one can view the experiment of P&S as having a partial 2x2 design in which they manipulate both the type of asset (single or contingent claims) and dividend values (homogeneous or heterogeneous). In Series A markets, different trader types had different dividend values for each state of nature, and there was a single asset trading in the market. In Series B markets, as in Series A, different types of traders had different dividend values, but there was a complete set of contingent claims, each of which only delivered a payout if the

⁹ Our study thus falls in the most general category of replication which uses a new research design to further test the hypothesis of the original study (Cartwright, 1991; Hunter, 2001; Maniadis, Tufano and List, 2014).

¹⁰ Surveys of experimental asset markets by Sunder (1995), Plott (2000), Deck and Porter (2013), and Morone and Nuzzo (2016) do not mention any such studies.

associated state was realized.¹¹ Thus, the difference between Series A and Series B is the type of asset(s) – single or contingent claim – that was tradable. In Series C, dividend values were the same for all traders, and there was a single tradable asset. The difference between Series A and Series C is the diverse or common dividend value of the single asset. Series A is arguably the most complex setup of the three; while Series B and Series C each introduce simplifications expected to facilitate information aggregation.

In P&S, the information aggregation hypothesis was rejected whenever the data on prices, allocation and profits were not better explained by the Rational Expectations model (hereafter RE) compared to two alternative models: Prior Information (i.e., Walrasian) and Maximin.¹² We briefly describe these models below.

Rational expectations (RE)

If the pooled information of all traders identifies the asset value with certainty, the fully-revealing rational expectations equilibrium price is the true asset value. To illustrate the mechanics, consider the case in which it is common knowledge that the asset value is 50, 240 or 490 each with equal probability. Suppose the true value is 50. In this case, half the traders would be given the signal “Not 240” and the other half would be given the signal “Not 490”. Before any trading occurs, the traders who receive the “Not 240” signal value the asset at 270; whereas, the “Not 490” traders value the asset at 145. These traders would also recognize that should anyone have received a “Not 50” signal that trader would value the asset at 365. The lack of traders willing to purchase the asset at a price between 270 and 365 thus reveals that none of the traders received the “Not 50” signal. This reveals that the true value must be 50, and so 50 would be the equilibrium price.

Prior information (PI)

In the prior information model, traders do not infer information regarding other traders’ signals from market prices, but they do apply Bayes’ rule to compute the expected asset value given the public information on the prior probability that a state is realized and their own private information (Lintner, 1969). Each individual trader is assumed to be willing to spend all her cash to buy assets at a price below her prior belief of its value and would sell all her certificates at any price above her prior belief. With sufficient liquidity, the market price will equal the prior belief of the trader type with the highest prior belief. Again using the example where 50, 240 and 490 are equally likely, a trader who received the signal “Not 50” would buy at prices below 365 and sell at prices higher than this. A trader who received the signal “Not 490” would

¹¹ Heterogeneous dividend values ensure that there are always gains from trade.

¹² We use the terminology of P&S to facilitate comparison of results thus referring to the Walrasian Model as Prior Information (PI).

buy at prices below 145 and sell at prices higher than this. Thus, if the true value is 240, then this would result in a price of 365.

Maximin (MM)

Similar to PI, market prices do not reveal all available private information in this model. Instead, a trader is assumed to buy the asset only when she is certain that the price is equal to or below the minimum dividend it could possibly yield, given the public information and her private information. With sufficient liquidity, the market price predicted by this model is equal to the minimum belief of the trader type with the highest belief of the minimum possible valuation of the asset. Returning to the case in which the true value is equally likely to be 50, 240 or 490, a trader who received the signal “Not 50” would buy at any price below 240 while a trader who received the signal “Not 240” would pay at most 50. Thus, if the true value is 490, then this would result in a price of 240.

In Series A, P&S rejected the hypothesis that market prices aggregated private information. However, P&S reported successful information aggregation in Series B and C when the RE model outperformed the PI and MM models (in pairwise tests) in predicting prices (see Table III in P&S, p. 1105), certificate allocations (Table IV in P&S, p. 1106) and profit distributions (Table V in P&S, p. 1107). These significant results are captured in Conclusions 2, 3 and 4 in P&S.

The authors note “The fact that both the second and third series perform as predicted by the RE model leads us to suspect that the existence of instruments which enable traders to link the actions of others to a source of motivation is important to the information aggregation function of markets” (P&S, p. 1086). Thus, P&S suggest that information aggregation is not a robust property of markets, but is possible under certain circumstances.

3. Replication Experiments

Following the replication procedures used by Camerer et al., 2016 (p. 1433), we attempted to replicate the most important statistically significant finding of the original study. For P&S that finding is that information aggregation successfully occurs in Series B and C markets. Our replication thus seeks to assess whether asset prices reveal all available private information in Series B and C markets. Following P&S, we compare the predictive ability of the RE model to that of PI and MM. Our replication attempt was preregistered in the AEA RCT Registry as AEARCTR-0003660.

Market Environment

The parameters for the markets are shown in Table 1, which is a subset of Table I (P&S, p. 1088). We replicated Markets 4, 5, 7, 8 and 9 which correspond to Series B and C. The only change from P&S is the exchange rate for converting lab earnings into cash payments. To

account for the cost of living difference in the approximately thirty years that have transpired since the original data were collected, the dollar per franc exchange rates have been doubled.¹³ Within each series, all traders in a session started with the same endowment of cash and certificates at the beginning of each market period. As the initial endowment of Francs represents a loan from the experimenter, traders had to pay a fixed cost at the end of each market to repay this loan, meaning that subjects' earnings were driven solely by dividends and capital gains/losses. The predicted prices under RE, PI and MM as well as the predicted ending allocation of certificates can be found in Table II (P&S, pp. 1091-1092) and are reproduced in Appendix A for each series, trader type, and realized state of nature.

Table 1. Market Parameters

Market (Series)	Trader Type	Number of Traders	Initial Endowment		Fixed Cost	\$/ Franc ^B	Dividends			Probabilities		
			Certificates	Francs			X	Y	Z	X	Y	Z
4 (B)	I	4	2	10,000	10,000	0.006	70	130	300	0.35	0.20	0.45
	II	4	2	10,000	10,000	0.006	230	90	60			
	III	4	2	10,000	10,000	0.006	100	160	200			
5 (B)	I	4	2	15,000	15,000	0.005	140	260	600	1/3	1/3	1/3
	II	4	2	15,000	15,000	0.005	460	180	120			
	III	4	2	15,000	15,000	0.005	200	320	400			
7 (C)	I	12	4	25,000	25,000	0.003	50	240	490	1/3	1/3	1/3
8 (C)	I	12	2	25,000	25,000	0.003	125	375	525	1/3	1/3	1/3
9 (C)	I	12	4	25,000	25,000	0.003	50	240	490	0.35	0.45	0.20

The exchange rate reported in Table I of P&S (p. 1088) shows the exchange rate for Market 9 was 0.015. We believe this was a typo as otherwise the payoffs for that market are not aligned with the payoffs for the other markets.

Power and Sample Size

Following Camerer et al., (2016), our aim was to ensure a statistical power of at least 90% to detect the P&S findings at a 5% significance level; thus, we adopt their method and estimate the multiple of the original sample size that would be needed to achieve 90% power by using the standard z-test power formula $\left(\frac{3.242}{z}\right)^2$ where z is the standard normal test statistic associated with the p-value of the original study. The main statistically significant results of P&S correspond to the comparison between the RE model and the other two models regarding prices (P&S, Table III), allocations (P&S, Table IV) and profits (P&S, Table V). Regarding the results on prices and profits, the required sample size is 3.019 times as large as the original sample, as the reported p-values are all equal to 0.031, which is associated with a z statistic of

¹³ The cost of living has indeed doubled in the US between 1988 and 2018 given an average annual inflation rate of 2.5% over that period. Note that the original experiments and our replication use an experimental currency called Francs that is converted into local currency at the end of the session. For most of P&S and all of our markets, that currency is dollars. However, Markets 7 and 8 of P&S were conducted in India and thus payments were converted into rupees.

1.866. Regarding allocations, the p-value is 0.048 so that the multiple of the original sample size that would be needed is equal to 3.791. We thus opted for a multiple of 4 and collected four times as much data as the original study in order to have sufficient statistical power. This value falls within the normal range of multiples used in recent replication attempts in economics and social sciences (see Camerer et al., 2016; 2018). We thus recruited a total of 144 subjects.¹⁴ All participants were recruited from Chapman University's Economic Science Institute subject pool and were undergraduate students at that institution with no prior experience in information aggregation markets.

Subject Experience and Earnings

Because the duration and required experience varied across the markets in P&S, subjects were recruited for different durations with different participation payments and average earnings.¹⁵ We were also careful to match the experience levels in P&S (see Table I, p. 1088). In Market 8, the subjects were experienced in Market 7. In P&S this means that the subjects were participating with the exact same group of people in the two markets, since there were 12 people in each market, and there was only one replicate of each market. However, in P&S the subjects in Market 5 were experienced either in Market 3 or 4. Thus, our Market 5 sessions were formed by combining half the subjects in one Market 4 session with half the subjects in another Market 4 session. The two remaining halves were also combined into a separate Market 5 session. Table 2 summarizes the details of our experimental sessions.

Procedures

Conduct of each session closely followed the procedures laid out in P&S. P&S (p. 1118) note that steps "were essentially similar to those described ... in Plott and Sunder (1982) with appropriate modifications to the instructions..." Thus, all handouts and instructions for this hand-run experiment are based on details in Plott and Sunder (1982) and reflect differences that were noted in P&S to the fullest degree possible. Each session consisted of five stages after subjects provided informed consent and a general script was read aloud. A description of the five stages as well as copies of instructions and handouts is available in Appendix B.

Finally, to guard against any bias from the auctioneer, the person serving in that role was unaware of the rational expectations price in every case and unfamiliar with the specifics of P&S, although that person had experience conducting double auctions in a classroom setting.

¹⁴ The number of subjects is less than 4 replicates x 5 markets x 12 traders because Markets 5 and 8 involved experienced traders as discussed in the next subsection.

¹⁵ The lab's policy required that subjects be compensated in part based upon study duration.

Table 2. Summary of Experimental Sessions

P&S Market and Series	Experience	Subjects / Session	Periods ^A	Sessions	Participation Payment ^B	Duration ^C	Average Earnings (USD) ^D
4 B	None	12	13	4	\$7	210 min	27.19
5 B	Market 4	12	9	4	\$30	150 min	33.89
7 C	None	12	14	4	\$7	180 min	39.78
8 C	Market 7	12	7	4	\$27	120 min	15.64
9 C	None	12	17	4	\$15	210 min	53.68

^A In periods 10-13 of Market 4 and 10-16 of Market 5, traders also completed Series A markets in P&S. We followed this procedure for Market 4, although we are not interested in analyzing Series A for which P&S did not report successful information aggregation. Throughout the remainder of this paper, when we refer to Market 4 we mean the Series B portion of the market in periods 1-9. We did not complete periods 10-16 (Series A markets) in Market 5 as, unlike Market 4, traders who participated in Market 5 are not recruited for other sessions. This means any additional experience of Series A markets at the end of Market 5 could not have impacted any behavior relevant to our replication endeavor.

^B The participation payments for sessions of Markets 5 and 8 were large to ensure that subjects returned from Markets 4 and 7 to ensure that we were able to match the experience levels of subjects in P&S (see Table I of P&S, p. 1088). Markets 7 and 8 were conducted on the same day, with a one-hour lunch break in between. Markets 4 and 5 were conducted either 1 or 2 days apart. For every Market 5 session, two subjects with experience in Market 9 were recruited as standbys in case a Market 4 subject did not return for Market 5. In one Market 5 session a single standby was used as an alternate. To further encourage the subjects to return for the second session, Market 4(7) earnings were not paid until the end of Market 5(8), with subjects only receiving the participation payment at the end of Market 4(7).

^C The duration denotes the length of time for which the participants were recruited.

^D The earnings exclude the participation payments.

4. Replication Results

In this section we present our results in a manner parallel to the presentation in P&S. For brevity, we show a graph (Figure 1) representing all asset prices in each period of the four replications of Market 9, the market most commonly adapted for subsequent research (Biais, Hilton, Mazurier and Pouget, 2005; Hanson, Oprea and Porter, 2006; Veiga and Vorsatz, 2010). Similar graphs for all the other markets can be found in Appendix C alongside additional statistical evidence. Horizontal lines indicate the predictions of the three competing models. The width of each period within a market is standardized to facilitate period-by-period comparisons across markets because despite the common duration of each period, the number of periods varies substantially by market.¹⁶ A first visual inspection suggests that prices differ from the RE model predictions regardless of the series and the market. In Table 3, we assess these differences using the same statistical approach as in P&S.

¹⁶ The trades are presented in sequential order, but, as in P&S, timing within the 7-minute trading period is not reflected.

Tables 3, C1, and C2 present the same analyses for the replication data that P&S present in Tables III, IV, and V for the original data (P&S, pp. 1105-1107), respectively. The data in these tables are used to assess Conclusions 2, 3, and 4 of P&S that RE will outperform both PI and MM models in predicting prices (Table 3), allocations (Table C1) and profits (Table C2). While P&S found support for RE over PI and MM, we did not replicate these conclusions, as the RE model did not outperform either model at a 5% significance level, regardless of whether we look at prices, allocations or profits.

Regarding prices, we assess the performance of the respective models using the three criteria proposed by P&S: mean absolute price deviation (hereafter MAD), log odds and percentage of convergent price changes. Following P&S, we give information aggregation its “best chance” by only considering the last occurrence of each of the possible asset values for each market (i.e., periods 7, 8 and 9 for Market 4 and periods 14, 15 and 17 for Market 9). To assess the MAD, we calculate the average across sessions of the absolute difference between the price and the value predicted by each model (PI, RE and MM) (see Appendix A or P&S). For each session, this value is calculated as:

$$MAD := \text{average}_{i,t} |p_{i,t} - m_t|$$

where i represents a transaction, t denotes a period in which a state was realized for the last time, $p_{i,t}$ corresponds to the transaction price, and m_t is the predicted price under the appropriate model. Thus, the mean absolute price deviation is computed as the average over all transactions in the last occurrence of each state of nature for each session. The second measure used in P&S is log odds, which we calculated by regressing, for each model, the observed prices in a given session on the predicted price under that model (using price data from the periods corresponding to the last occurrence of each state of nature). We then recovered the log likelihood value under the normality assumption. A third measure of information aggregation developed in P&S is the percentage of convergent price changes. The $N + 1^{st}$ transaction is considered convergent if its price is no further from the selected model’s prediction than the price of the previous transaction, N . That is, the $N + 1^{st}$ transaction is deemed convergent if:

$$|p_{N+1,t} - m_t| \leq |p_{N,t} - m_t|$$

where p_N is the market price of the N^{th} transaction. For each model, the ratio of convergent price changes to the total number of transactions over the three periods in which a state was realized for the last time per session is reported in Table 3.¹⁷

¹⁷ As an initial transaction is required to determine if the subsequent transaction is convergent, we utilize the total number of transactions minus three (since we consider three market periods) as the denominator in this ratio.

An overview of Table 3 shows that the RE model does not outperform the other two models, regardless of the criterion or the series we consider. The RE model actually underperforms the PI model on the criterion of mean absolute price deviation (p-value = 0.048) when considering, as in P&S, all sessions from Series B and C. But, the RE model does outperform PI on the criterion of percentage of convergent price changes even though RE does not significantly outperform MM for this criterion. No model stands out with regard to the criterion of log odds. If we take Series B and C separately, we obtain a similar finding; RE does not perform significantly better than both of the other two models according to any of the three criteria. This contradicts Conclusion 2 in P&S (p. 1104) according to which “In markets with a complete set of state contingent securities (Series B) and in markets with a single security with uniform dividends (Series C) the RE model price predictions outperform both PI and MM.”

Finding 1: We are unable to replicate the finding that the RE model outperforms the PI and MM models in terms of predicting prices in the Series B and C markets.

Here we report two additional findings indicating that our results also hold when looking at Allocations and Profits, with the supporting evidence presented in Tables C1 and C2 of Appendix C to conserve space.

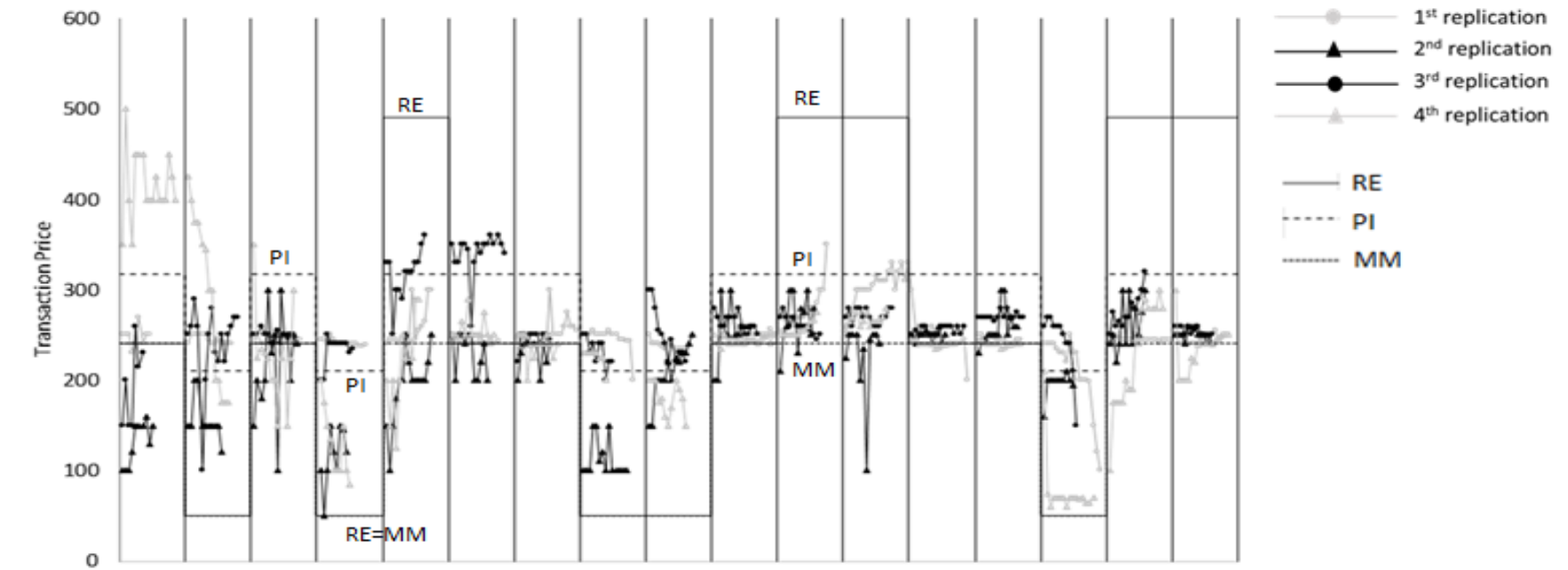
Finding 2: We are unable to replicate the finding that RE outperforms PI and MM in predicting final asset allocations in the Series B and C markets, though RE outperforms PI when evaluating Series B markets separately

Finding 3: We are unable to replicate the finding that RE outperforms PI and MM in predicting traders’ profits in the Series B and C markets, though RE outperforms PI and MM when evaluating Series C markets separately.

As a final note on our replication, experience does not seem to encourage information aggregation in our sample, as evidence for information aggregation was not found despite using for each session, as in P&S, only the data for the last occurrence of each possible state of nature.¹⁸ Moreover, we do not find support for the RE model prediction when restricting our analysis to the 8 sessions in which the market was populated exclusively by experienced traders (see Appendix D for the analysis of Markets 5 and 8 combined).

¹⁸ The general conclusions do not differ if data from every trading period is used.

Figure 1. Prices in Market 9 Sessions



Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
State	Y	X	Y	X	Z	Y	Y	X	X	Y	Z	Z	Y	Y	X	Z	Z

Average Prices

P&S	253	100	229	95	458	240	245	50	175	242	490	480	240	240	50	489	489
Rep	278	237	235	189	247	271	244	201	218	251	267	273	248	254	175	250	245

Table 3. Comparison of Actual Prices to Prices Predicted by Three Models at the End of Each Market

Series	Market Experiments		MAD			Log Odds Under Normality			Percentage of Convergent Price Changes		
	Market	Replicate	PI	RE	MM	PI	RE	MM	PI	RE	MM
B	4	1	78.30	100.38	90.88	-216.24	-225.24	-239.01	37.50	66.67	66.67
		2	81.30	97.20	58.58	-183.83	-184.21	-212.05	37.50	66.67	62.50
		3	56.10	109.00	89.92	-197.98	-210.01	-232.44	54.17	75.00	62.50
		4	50.61	89.80	93.86	-224.90	-250.29	-265.10	32.14	75.00	85.71
	5	1	58.10	205.00	215.90	-686.42	-677.50	-781.58	38.10	66.67	33.33
		2	109.00	159.00	252.20	-731.43	-686.31	-814.38	40.91	77.27	31.82
		3	95.70	143.00	281.50	-726.08	-691.72	-840.57	35.71	71.43	35.71
		4	95.86	182.00	257.30	-629.24	-612.03	-698.57	40.00	75.00	20.00
C	7	1	98.80	104.05	54.05	-188.57	-262.57	-250.43	56.00	60.00	44.00
		2	131.00	50.60	31.11	-206.06	-196.63	-186.21	53.57	50.00	46.43
		3	152.00	17.00	39.65	-402.91	-332.67	-372.38	47.62	66.67	33.33
		4	93.78	139.00	156.10	-272.60	-338.71	-318.74	46.88	50.00	53.13
	8	1	134.00	56.20	31.23	-150.20	-154.51	-142.17	53.85	46.15	69.23
		2	103.00	14.10	55.40	-182.75	-156.68	-170.40	43.75	50.00	56.25
		3	75.00	125.00	157.80	-183.25	-211.84	-209.95	36.36	45.45	36.36
		4	35.79	82.20	71.27	-158.08	-200.83	-194.25	50.00	35.00	40.00
9	1	57.20	145.00	61.53	-244.81	-306.81	-294.03	53.57	75.00	60.71	
	2	48.30	131.00	53.00	-154.65	-215.91	-201.88	50.00	42.86	57.14	
	3	50.60	143.00	72.50	-193.37	-241.66	-229.72	54.55	59.09	59.09	
	4	98.11	102.00	15.96	-249.64	-268.37	-226.30	56.00	68.00	60.00	
Wilcoxon Signed Ranked Sum Tests			↑	↑	↑	↑	↑	↑	↑	↑	↑
B-series Sessions			↓	↓		↓	↓		↓	↓	
C-series Sessions											
All Sessions											
			0	25		22	36		36	32	
			0.012(PI)	0.327(RE)		0.575(RE)	0.012(RE)		0.012 (RE)	0.042(RE)	
			35	19		61	59		25	47	
			0.754(PI)	0.117(MM)		0.084(PI)	0.117(MM)		0.272(PI)	0.505(RE)	
			61	96		140	144		195	154	
			0.100(PI)	0.654(MM)		0.191(PI)	0.145(RE)		0.003(RE)	0.059(RE)	

Reported metrics for each session are based on the last period in which the state was X, the last period in which the state was Y and the last period in which the state was Z. Each statistical test compares RE to an alternative model. Each cell in the lower portion of the table contains the associated test statistic and level of significance as well. The model in parentheses is the one favored by the data based on the median. This table is comparable to Table III in P&S.

5. Post-Study Probability that Series B and C Markets Aggregate Information

It is reasonable to ask what the balance of evidence suggests about the likelihood that Series B- and C-style markets behave in a manner more consistent with RE than PI or MM. This in part depends on one's initial belief about the probability of this pattern holding. Denoting one's prior as π , Maniadis, Tufano and List (2014) show that following a rejection of the null hypothesis of no effect in favor of the alternative, the PSP is given by:

$$PSP \text{ following a rejection} = \frac{(1-\beta)\pi}{(1-\beta)\pi + \alpha(1-\pi)} \quad (1)$$

where α and β denote the probability of type I and II errors respectively. If, on the other hand, one fails to reject the null in favor of the alternative then the PSP (of a treatment effect) is given by:

$$PSP \text{ following a failure to reject} = \frac{\beta\pi}{\beta\pi + (1-\alpha)(1-\pi)} \quad (2)$$

Focusing on MAD and treating the z-scores we used to calculate the desired sample size for our replication as our measure of the effect sizes (i.e. of the average difference between MAD under RE and MAD under an alternative model), one can determine the ex-post power of P&S to detect that effect size via the Wilcoxon signed-rank sum test.¹⁹ For the comparison to PI their power (equal to $1 - \beta$) was 0.9457 and for comparison to MM it was 0.8963 using the conventional significance level of 0.05. Table 4 shows the relationship between prior beliefs and posterior beliefs based on the evidence put forward by P&S using equation (1). As is evidenced from the table, the PSP is quite high unless one held an extremely low prior. Because of our replication sessions, P&S is not the only evidence available. Maintaining the assumption that the true effect size is as large as is implied by the p-value reported in P&S, our replication had a power greater than 0.9999 both for the comparison of RE and PI as well as for the comparison of RE and MM under the Wilcoxon signed-rank sum test. Applying equation (2) to the resultant posterior beliefs after observing the evidence from P&S yields the PSP for the combined evidence from the two sets of experiments. As is clear from Table 4, the new evidence is enough to sway beliefs dramatically in the opposite direction for all but those who originally believed RE outperformed other models with near certainty.

¹⁹ Ioannidis (2008) and Loken and Gelman (2017) have argued that reported effects in small-sample studies are likely biased upwards because of the "statistical significance filter" (i.e. the fact that a threshold is used to select which results are publishable and which are not, and to meet that threshold in a small sample study, which will typically have large standard errors, requires a large effect). However, we assume that the treatment effect observed by P&S is the true effect size as assuming the true effect size is smaller would reduce the power associated with their test. In Appendix F, we report additional analysis of PSP assuming the true effect size is two thirds as large and show that our results are qualitatively unchanged.

Table 4. Post Study Probability that MAD Lower Under RE than PI or MM

<u>Probability RE outperforms PI</u>			<u>Probability RE outperforms MM</u>		
Prior	PSP after P&S	PSP after Replication	Prior	PSP after P&S	PSP after Replication
0.99	0.9995	0.1646	0.99	0.9994	0.1574
0.95	0.9972	0.0364	0.95	0.9971	0.0346
0.9	0.9942	0.0176	0.9	0.9938	0.0167
0.75	0.9827	0.0059	0.75	0.9817	0.0056
0.5	0.9498	0.0020	0.5	0.9472	0.0019
0.25	0.8631	0.0007	0.25	0.8566	0.0006
0.1	0.6776	0.0002	0.1	0.6658	0.0002
0.05	0.4989	0.0001	0.05	0.4855	0.0001
0.01	0.1604	<0.0001	0.01	0.1533	<0.0001

6. New Series D Markets

With a goal of showing that information aggregation *could* happen, it would be reasonable for P&S to stop the search after observing their positive results. However, with so few independent markets, P&S offer little evidence as to the robustness of their findings. Because our study failed to replicate the original findings, the boundary conditions for markets to aggregate dispersed information must be redrawn. We thus extended the P&S design to identify sufficient conditions for information aggregation with the aim of providing useful insights to practitioners.

The failure of the *efficient market hypothesis* is at odds with the empirical evidence from field data suggesting that prediction markets are effective in aggregating information as discussed in the introduction. Careful inspection of functioning prediction markets in practice reveals that they do not resemble Series B- or Series C-style markets. Instead, these markets are typically structured as contingent claims markets with homogeneous values.²⁰ That is, in practice markets combine both institutional features that were introduced separately by P&S to help facilitate efficient markets. Therefore, we ask if this combination will lead to successful information aggregation. That is, we complete the 2×2 experimental design of P&S by introducing Series D markets.

²⁰ Series A markets with a single asset and heterogeneous values can be thought of as markets for real assets. For example, an umbrella will have a high value to a person wearing suede when it is raining, a fair skinned person when it is sunny, and no one when it is overcast. Stocks and most other financial instruments are best described as Series C markets.

We introduce two new markets, which we refer to as Markets 12 and 13 to avoid confusion with the market numbering used in P&S. Table 5 provides the relevant market parameters, which are adapted from those in P&S. Predictions for prices and allocations for these markets can be found in Appendix A.

Table 5. Market Parameters

Market (Series)	Trader Type	Number of Traders	Initial Endowment		Fixed Cost	\$/ Franc ^B	Dividends			Probabilities		
			Certificates	Francs			X	Y	Z	X	Y	Z
12 (D)	I	12	2	25,000	25,000	0.006	125	375	525	1/3	1/3	1/3
13 (D)	I	12	2	25,000	25,000	0.006	50	240	490	0.35	0.45	0.20

The experimental procedures followed those used for the replication of Series B and C markets. Table 6 provides details for the experimental sessions. The subjects who participated in the Series D markets were drawn from the same pool as those who participated in the replication study, but none had participated in those sessions or any other related studies.

Table 6. Summary of Experimental Sessions

Market and Series	Experience	Subjects/ Session	Periods	Sessions	Participation Payment	Duration ^A	Average Earnings (USD) ^B
12 D	None	12	13	4	\$10	180 min	52.99
13 D	None	12	13	4	\$10	180 min	49.36

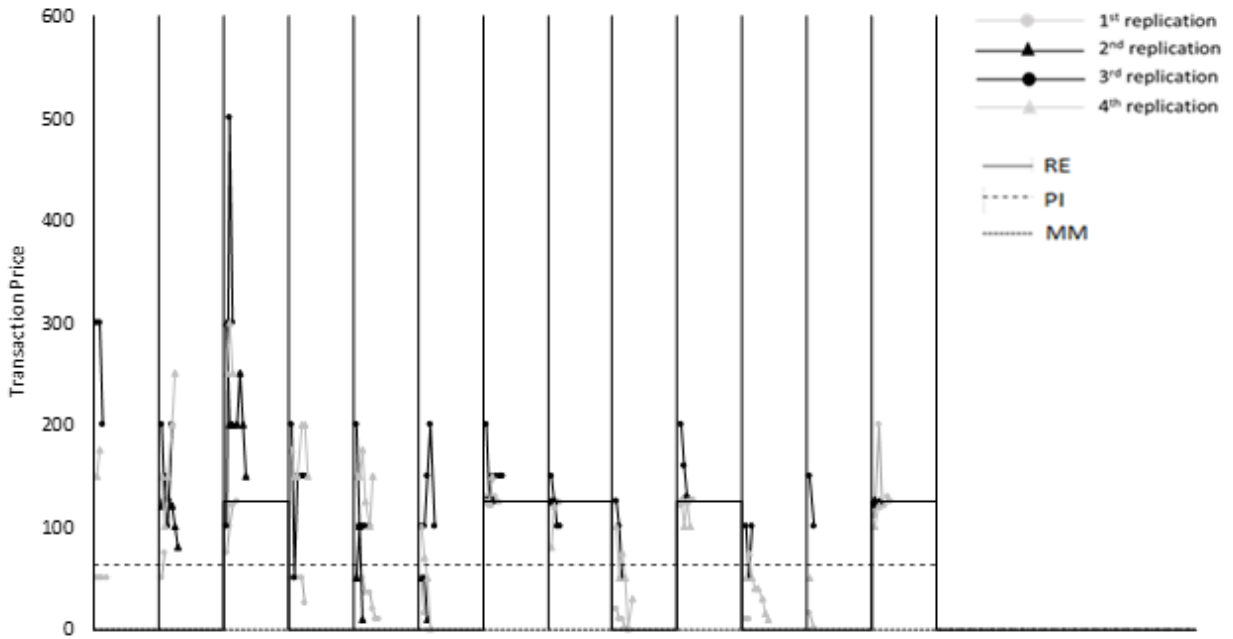
^A The duration denotes the length of time for which the participants were recruited.

^B The earnings exclude the participation payments.

Figures 2 and 3 show the transaction prices by period in each session for Markets 12 and 13, respectively, as well as the RE, PI and MM predictions. The results are dramatically different from what we observed in our replication of Series B and C markets. For Series D, prices consistently follow the RE predictions, reminiscent of the stark figures in P&S.

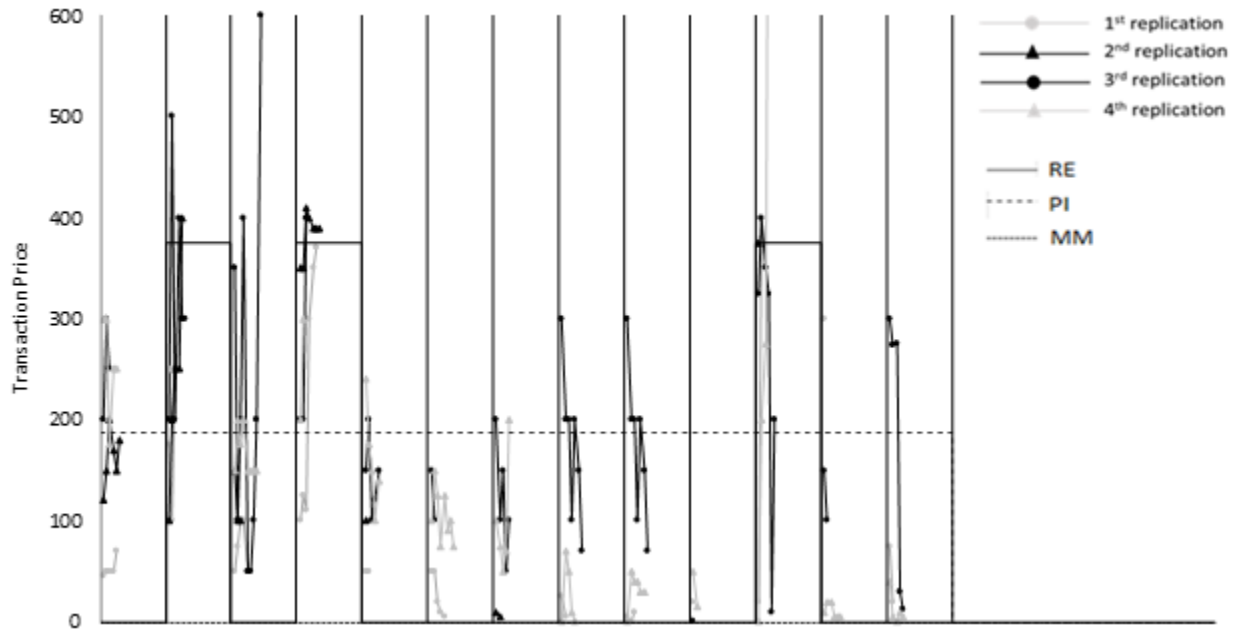
For the statistical analysis we rely on the same series of tests employed for Series B and C. The one caveat is that we do not analyze the allocation of assets for Series D. This is because with homogeneous values, the rational expectations model makes no prediction regarding who should hold the assets after trading is completed. We also note that RE and MM make the same profit predictions for Series D markets: both predict each subject will have the average earnings. Thus, profit can be used to compare PI to RE and PI to MM, but cannot be used to differentiate between RE and MM.

Figure 2a. Prices in Market 12 Sessions – Asset X



Period	1	2	3	4	5	6	7	8	9	10	11	12	13
State	Z	Y	X	Y	Z	Z	X	X	Z	X	Y	Z	X
<u>Average Prices</u>	147	135	213	122	85	75	137	119	48	130	45	63	127

Figure 2b. Prices in Market 12 Sessions – Asset Y

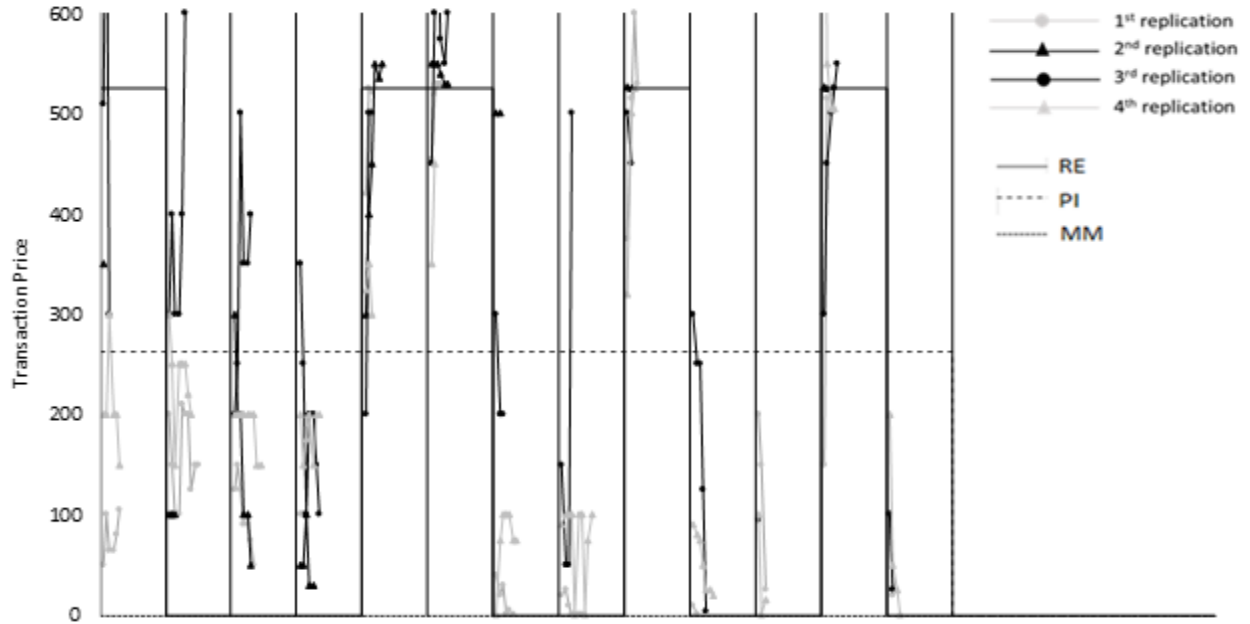


Period	1	2	3	4	5	6	7	8	9	10	11	12	13
State	Z	Y	X	Y	Z	Z	X	X	Z	X	Y	Z	X

Average Prices

	172	258	164	296	133	82	93	99	89	22	297	68	87
--	-----	-----	-----	-----	-----	----	----	----	----	----	-----	----	----

Figure 2c. Prices in Market 12 Sessions – Asset Z

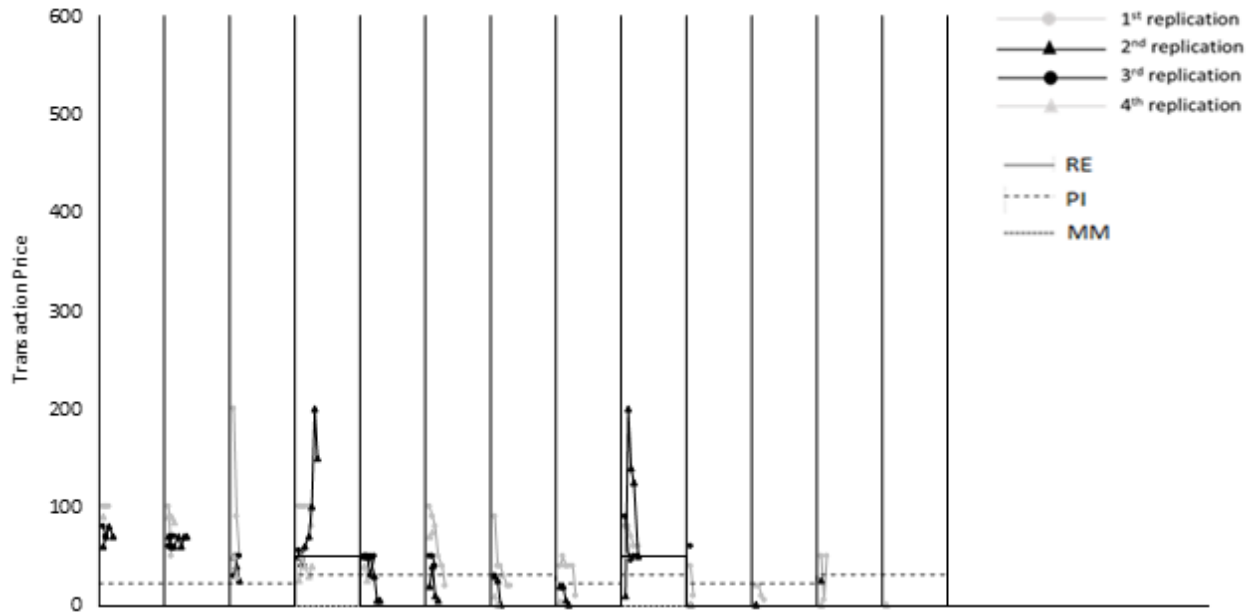


Period	1	2	3	4	5	6	7	8	9	10	11	12	13
State	Z	Y	X	Y	Z	Z	X	X	Z	X	Y	Z	X

Average Prices

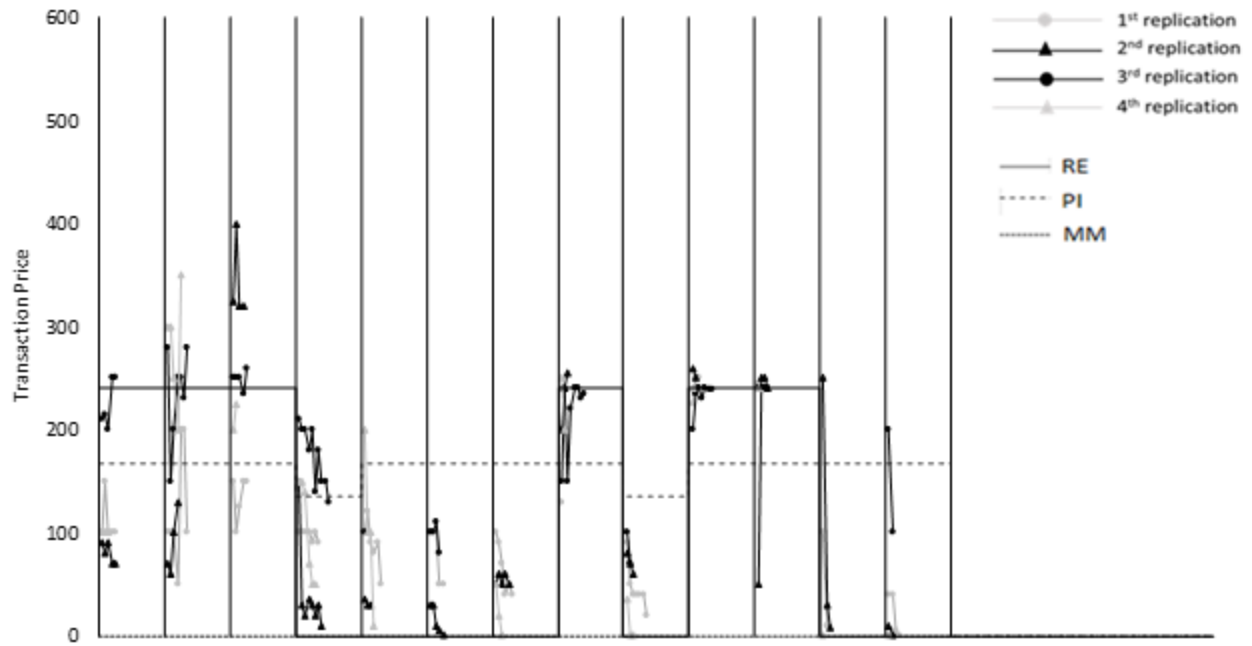
230	224	193	147	428	531	129	75	488	93	84	510	65
-----	-----	-----	-----	-----	-----	-----	----	-----	----	----	-----	----

Figure 3a. Prices in Market 13 Sessions – Asset X



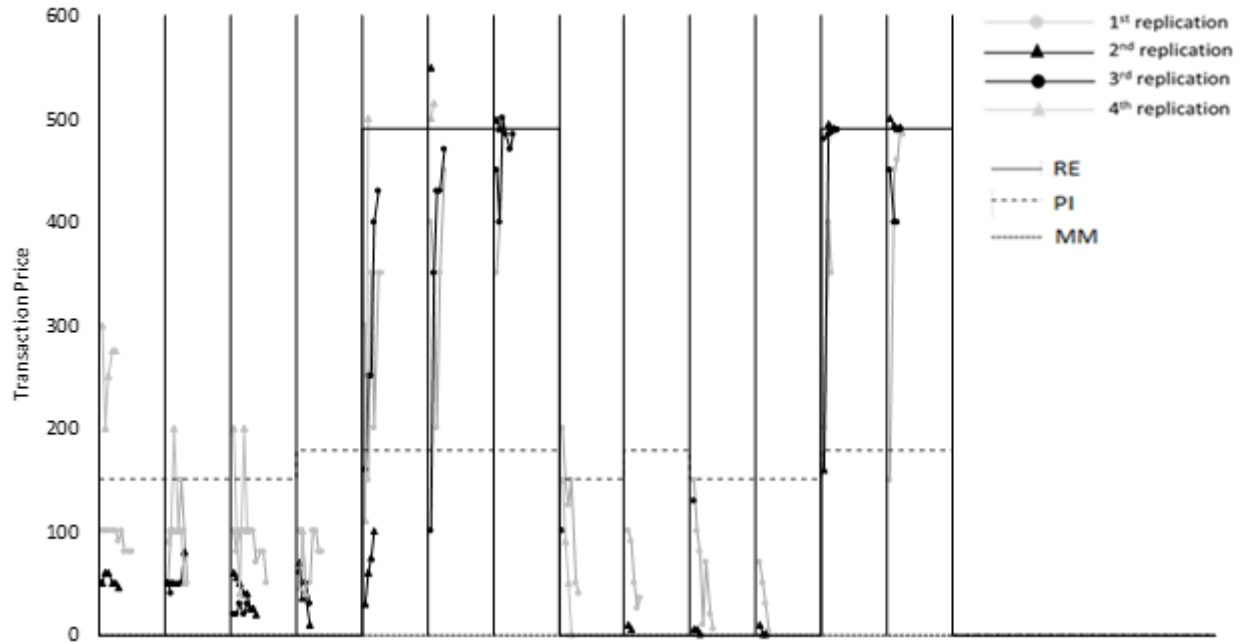
Period	1	2	3	4	5	6	7	8	9	10	11	12	13
State	Y	Y	Y	X	Z	Z	Z	Y	X	Y	Y	Z	Z
<u>Average Prices</u>	82	72	60	74	40	49	28	25	76	22	11	26	1

Figure 3b. Prices in Market 13 Sessions – Asset Y



Period	1	2	3	4	5	6	7	8	9	10	11	12	13
State	Y	Y	Y	X	Z	Z	Z	Y	X	Y	Y	Z	Z
<u>Average Prices</u>	132	185	232	108	81	62	52	213	49	238	224	67	44

Figure 3c. Prices in Market 13 Sessions – Asset Z



Period	1	2	3	4	5	6	7	8	9	10	11	12	13
State	Y	Y	Y	X	Z	Z	Z	Y	X	Y	Y	Z	Z

Average Prices

118	81	70	63	239	392	453	101	45	52	24	418	439
-----	----	----	----	-----	-----	-----	-----	----	----	----	-----	-----

Table 7. Comparison of Actual Prices to Prices Predicted by Three Models at the End of Each Market

<u>Market Experiments</u>			<u>MAD</u>			<u>Log Odds Under Normality</u>			<u>Percentage of Convergent Price Changes</u>		
Series	Market	Replicate	PI	RE	MM	PI	RE	MM	PI	RE	MM
D	12	1	129.38	74.04	160.29	-950.79	-914.53	-998.76	21.43	78.57	57.14
		2	168.00	0.40	335.40	-542.71	-525.96	-597.05	100	100	100
		3	127.00	108.00	219.10	-1029.02	-1077.74	-1135.40	33.33	66.67	55.56
		4	170.70	73.50	138.70	-1191.32	-1198.80	-1258.64	26.09	78.26	65.22
	13	1	107.08	41.03	101.31	-1113.53	-1083.91	-1200.78	20.00	90.00	80.00
		2	140.00	34.60	147.50	-901.13	-852.67	-940.53	50.00	85.71	78.57
		3	83.50	45.94	185.30	-724.62	-691.81	-813.43	50.00	66.67	83.33
		4	93.86	5.71	47.14	-490.26	-494.68	-530.05	0.00	100.00	100.00
Wilcoxon Signed Ranked Sum Tests			↑ _____ ↑	↑ _____ ↑		↑ _____ ↑	↑ _____ ↑		↑ _____ ↑	↑ _____ ↑	
D-series Sessions			↓	↓		↓	↓		↓	↓	
B&D Sessions			36	36		25	36		35	26	
			0.012 (RE)	0.012 (RE)		0.327 (RE)	0.012 (RE)		0.014 (RE)	0.178 (RE)	
C&D Sessions			80	125		94	136		135	112	
			0.535(RE)	0.003(RE)		0.178(RE)	<0.001(RE)		<0.001(RE)	0.015(RE)	
All Sessions			148	138		79	150		180	137	
			0.108(RE)	0.218(RE)		0.332(PI)	0.093(RE)		0.005(RE)	0.190(RE)	
			218	276		218	339		376	298	
			0.733(RE)	0.096(RE)		0.585(PI)	0.002(RE)		<0.001(RE)	0.022(RE)	

Reported metrics for each session are based on the last period in which the state was X, the last period in which the state was Y and the last period in which the state was Z. Each statistical test compares RE to an alternative model. Each cell in the lower portion of the table contains the associated test statistic and level of significance as well. The model in parentheses is the one favored by the data based on the median. This table is comparable to Table III in P&S.

The statistical results are reported in Tables 7 and E1 (in appendix E). The analysis strongly supports RE over both PI and MM. The MAD is lower under RE than PI or MM in all eight sessions (p-values < 0.001).

Combining the Series D data with the data from Series B and C, we can compare models for all contingent claims markets (Series B and D), homogeneous value markets (Series C and D), and all markets (Series B, C and D). These tests can be found in the lower portions of Tables 7 and E1. The statistical results favor the RE model over either of the other two models to varying levels of significance. Only the markets which feature both contingent claims and homogeneous values (Series D) provide consistent support for the RE model across the various tests.

So how convinced should one be that Series D markets aggregate information? Table 8 provides the PSP using equation (1) for various priors that one might hold, assuming the actual MAD treatment effect equals the z-score associated with our reported p-value of 0.012 (note that this is the p-value for comparisons of RE to each of the other two models; we use the z-score for consistency with our PSP analysis of Series B and C markets). Under this assumption, our ex post power for the comparison between RE and PI and for the comparison between RE and MM was greater than 0.9999 given a significance level of 0.05.²¹ While the implication from the table is that all but the very skeptical should conclude that Series D markets do aggregate information, it is worth noting again that a similar conclusion would have been drawn for Series B and C markets based on the data in P&S alone. Thus, our own replication of P&S calls for caution with regards to our Series D results until independent replication of our findings is conducted.

Table 8. Post Study Probability that MAD Lower Under RE than PI or MM for Series D

<u>Probability RE outperforms PI</u>		<u>Probability RE outperforms MM</u>	
Prior	PSP	Prior	PSP
0.99	0.9995	0.99	0.9995
0.95	0.9974	0.95	0.9974
0.9	0.9945	0.9	0.9945
0.75	0.9836	0.75	0.9836
0.5	0.9524	0.5	0.9524
0.25	0.8696	0.25	0.8696
0.1	0.6896	0.1	0.6896
0.05	0.5128	0.05	0.5128
0.01	0.1681	0.01	0.1681

²¹ As with Series B and C, if we assume that the true effect size is only two thirds of the reported effect size, the results remain qualitatively unchanged. See Table F2 in Appendix F.

7. Discussion

Whether markets are able to aggregate dispersed information is a fundamental question in Economics (Hayek, 1945). Because testing the strong-form efficiency of markets is not possible using archival data (Fama, 1991), experimental markets are critical to test whether such efficiency is achievable. The breakthrough design proposed by P&S thus made the *efficient market hypothesis* falsifiable. Although the original findings in P&S show that it is possible for markets to achieve strong-form efficiency, our data show these findings are even more fragile than P&S recognized. Moreover, neither the existence of contingent claims nor common dividend values alone are sufficient to facilitate robust information aggregation.

Completing the experimental design of P&S by introducing Series D markets, which incorporate both contingent claims and homogeneous values, yields broad support for the *efficient market hypothesis*. Further, this market structure is more similar to prediction markets operating in practice than any of the three markets considered by P&S, which suggests that market designers may have discovered this combination despite the absence of a direct institutional comparison in the literature. This finding provides a key insight regarding institutional design for those assessing and designing markets in the lab or in the field. A crucial implication of our findings is that a complete set of contingent claims will not guarantee strong-form efficiency. This means that carefully-designed markets might still be inefficient if the market is populated by agents with different preferences over assets. Interestingly, these differences are likely to be magnified by the increasing globalization of financial markets because traders will more often face different regulatory constraints and tax systems and will more likely exhibit different attitudes toward risk (Kose, Prasad, Rogoff and Wei, 2009). It follows that the informational efficiency of markets might decrease as a result of the increased diversity of market participants.

While our work fails to replicate the results of P&S in a direct sense, our findings do support the broader notion that markets can aggregate information. While their findings showed it is possible for markets to aggregate information, our results reveal conditions under which such aggregation is robust. As described by P&S, both contingent claims and homogeneous values are features that provide “some knowledge about other traders’ preferences” (p. 1117).²² The ability to infer the intentions of other traders from their market actions and then update one’s own private information accordingly is critical for efficient markets. Thus, the level of traders’ sophistication is likely a key component of market performance. Hence, the one feature of P&S

²² It is also the case that the contingent claims effectively introduce insiders in all but one of the markets. Specifically, for every state that does not occur some traders know that the associated asset is worth 0 (e.g. if the true state is X, then consider the markets for Y and Z). Plott and Sunder (1982) found that markets with fully informed insiders had prices that reflected that information. Our Series B results are evidence that insiders are not sufficient for information dissemination.

that we did not replicate, namely the subject pool, may be partially driving some of the observed differences.²³ Recent work by Corgnet, DeSantis and Porter (2015) supports this conclusion as they find that markets with more cognitively sophisticated traders (as measured with cognitive reflection, Frederick, 2005) are more likely to have prices consistent with the rational expectations hypothesis. Such a pattern again highlights the importance of effective market design. The lack of robustness of information aggregation results contrasts with numerous findings showing the robustness of competitive equilibria outcomes to various institutions (Smith, 1962; List, 2004) and levels of traders' sophistication (Becker, 1962; Gode and Sunder, 1993). Following Radner (1982), we posit that rational-expectation models are particularly demanding regarding the sophistication of economic agents which might explain the lack of robustness of their predictions compared to competitive equilibria predictions that do not hinge on rational expectations.

²³ P&S collected data from student subjects at Caltech and the Indian Institute of Management.

8. References

- Anderson, L. and C. Holt (1997) "Information Cascades in the Laboratory," *American Economic Review*, 847-862.
- Arrow, Kenneth, Robert Forsythe, Michael Gorham, Robert Hahn, Robin Hanson, John Ledyard, Saul Levmore, Robert Litan, Paul Milgrom, Forrest Nelson, George Neumann, Marco Ottaviani, Thomas Schelling, Robert Shiller, Vernon Smith, Erik Snowberg, Cass Sunstein, Paul Tetlock, Philip Tetlock, Hal Varian, Justin Wolfers, and Eric Zitzewitz (2008): "The Promise of Prediction Markets," *Science*, 320, 877-888.
- Atanasov, P., P. Rescober, E. Stone, S. A. Swift, E. Servan-Schreiber, P. Tetlock, L. Ungar, and B. Mellers (2017) "Distilling the Wisdom of Crowds: Prediction Markets vs. Prediction Polls," *Management Science* 63(3), 691-706.
- Becker, G. S. (1962). "Irrational Behavior and Economic Theory." *Journal of Political Economy* 70(1), 1-13.
- Berg J., R. Forsythe, F. Nelson, T. Rietz (2001) "Results From a Dozen Years of Election Futures Markets Research." in Plott CR, Smith VL, eds. *Handbook of Experimental Economics Results*, Vol. 1, NorthHolland, Amsterdam, 742–751.
- Biais, Bruno, Denis Hilton, Karine Mazurier and Sébastien Pouget (2005) "Judgmental Overconfidence, Self-Monitoring and Trading Performance in an Experimental Financial Market," *The Review of Economic Studies* 72(2), 287-312.
- Bossaerts, P. (2009) "What Decision Neuroscience Teaches us About Financial Decision Making," *Annual Review of Financial Economics* 1, 383-404.
- Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler (1997). "Labor Supply of New York City Cabdrivers: One Day at a Time," *The Quarterly Journal of Economics* 112(2), 407-441.
- Camerer C., Dreber A., E. Forsell, T.H. Ho, J. Huber, M. Johannesson, M. Kirchler, J. Almenberg, A. Altmejd, T. Chan, E. Heikensten, F. Holzmeister, T. Imai, S. Isaksson, G. Nave, T. Pfeiffer, M. Raza, and H. Wu (2016) "Evaluating Replicability of Laboratory Experiments in Economics," *Science* 351, 1433-1436.
- Camerer C., Dreber A., F. Holzmeister, T.-H. Ho, J. Huber, M. Johannesson, M. Kirchsler, G. Nave, B. Nosek, T. Pfeiffer, A. Altmejd, N. Buttrick, T. Chan, Y. Chen, E. Forsell, A. Gampa, E. Heikensten, L. Hummer, T. Imai, S. Isaksson, D. Manfredi, J. Rose, E.-J. Wagenmakers, and H. Wu (2018) "Evaluating Replicability of Social Science Experiments in Nature and Science between 2010 and 2015," *Nature Human Behaviour* 2, 637-644.
- Cartwright, N. (1991) "Replicability, Reproducibility, and Robustness: Comments on Harry Collins," *History of Political Economy* 23(1), 143-155.
- Chamberlin, E. (1948) "An Experimental Imperfect Market" *The Journal of Political Economy* 56, 95-108.

Chen, K. and C. Plott (2002) "Information Aggregation Mechanisms, Concept, Design and Field Implementation," Social Science Working Paper 1131, California Institute of Technology.

Choo, L., T. Kaplan, and R. Zultan (2019) "Information Aggregation in Arrow-Debreu Markets: An Experiment," *Experimental Economics* 22, 625-652.

Corgnet, B., Deck, C., DeSantis, M. and D. Porter (2018) "Information (Non) Aggregation in Markets with Costly Signal Acquisition," *Journal of Economic Behavior & Organization* 154, 286-320.

Corgnet, B., DeSantis, M. and D. Porter (2015) "Revisiting Information Aggregation in Asset Markets: Reflective Learning & Market Efficiency," ESI Working Paper 15-15.

Corgnet, B., DeSantis, M. and D. Porter (2019) "The Distribution of Information and the Price Efficiency of Markets", In Press *Journal of Economic Dynamics and Control*.

Court, D., B. Gillen, J. McKenzie and C. Plott (2018) "Two Information Aggregation Mechanisms for Predicting the Opening Weekend Box Office Revenues of Film: Boxoffice Prophecy and Guess of Guesses," *Economic Theory* 65, 25-54.

Cowgill, B., J. Wolfers, and E. Zitzewitz (2009) "Using Prediction Markets to Track Information Flows: Evidence from Google," in *Auctions, Market Mechanisms and Their Applications* Volume 14, eds. Sanmay Das, Michael Ostrovsky, David Pennock, and Boleslaw Szymanski, Springer: Berlin.

Cowgill, B. and E. Zitzewitz (2015) "Corporate Prediction Markets: Evidence from Google, Ford, and Firm X," *The Review of Economics Studies* 82(4), 1309-1341.

De Long, J.B. and K. Lang (1992). "Are All Economic Hypotheses False?" *Journal of Political Economy* 100(6), 1257-1272.

Deck, C. and D. Porter (2013) "Prediction Markets in the Laboratory," *Journal of Economic Surveys* 27(3), 589-603.

Fama, E. (1970) "Efficient Capital Markets: A Review of Theory and Empirical Work," *The Journal of Finance* 25(2), 383-417.

Fama, E. (1991) "Efficient Capital Markets: II," *The Journal of Finance* 46(5), 1575-1617.

Fama, E. (2008) "A Brief History of the Efficient Market Hypothesis," American Finance Association speech.

Forsythe, R. and R. Lundholm (1990) "Information Aggregation in an Experimental Market," *Econometrica* 58(2), 309-348.

- Frederick, S. (2005) "Cognitive Reflection and Decision Making," *Journal of Economic perspectives* 19(4), 25-42.
- Frydman, C., Barberis, N., Camerer, C., Bossaerts, P. and A. Rangel (2014) "Using Neural Data To Test a Theory of Investor Behavior: An Application to Realization Utility," *Journal of Finance* 69, 907-946.
- Gillen, B., C. Plott, and M. Shum (2017) "A Pari-Mutuel-Like Mechanism for Information Aggregation: A Field Test inside Intel," *Journal of Political Economy* 125(4), 1075-1099.
- Gode, D. K., and S. Sunder (1993) "Allocative Efficiency of Markets with Zero-intelligence Traders: Market as a Partial Substitute for Individual Rationality." *Journal of Political Economy* 101(1), 119-137.
- Gürkaynak, R. S. and J. Wolfers (2006) "Macroeconomic Derivatives: An Initial Analysis of Market-based Macro Forecasts, Uncertainty, and Risk. NBER Working Paper 11929.
- Hampton, K. (2007) "The Double-Auction gambling Market: An Experimental Examination," *The American Journal of Economics and Sociology* 66(3), 493-532.
- Hanson, R., Oprea, R. and D. Porter (2006) "Information Aggregation and Manipulation in an Experimental Market," *Journal of Economic Behavior and Organization* 60, 449-459.
- Hayek, F. (1945) "The Use of Knowledge in Society," *American Economic Review* 35(4), 519-530.
- Huber, J., M. Angerer and M. Kirchler (2011) "Experimental Asset Markets with Endogenous Choice of Costly Asymmetric Information," *Experimental Economics* 14(2), 223-240.
- Hunter, J. (2001) "The Desperate Need for Replications," *Journal of Consumer Research* 28(1), 149-158.
- Ioannidis, J. (2005) "Why Most Published Research Findings Are False," *PLoS Medicine* 2(8), 1418-1422.
- Ioannidis, J. (2008) "Why Most Discovered True Associations Are Inflated," *Epidemiology* 19(5), 640-648.
- Kose, M. A., E. Prasad, K., Rogoff and S. J. Wei (2009) "Financial Globalization: A Reappraisal," *IMF Staff papers* 56(1), 8-62.
- Kwasnica, A., R. Velthuis, and J. Williams (2019) "Did You See What I saw? Interpreting Others' Forecasts When Their Information is Unknown", *Review of Finance* 23(2), 325-361.
- Lintner, J. (1969) "The Aggregation of Investor's Diverse Judgments And Preferences In Purely Competitive Security Markets," *Journal of Financial and Quantitative Analysis* 4, 347-400.

- List, J. (2004). "Testing Neoclassical Competitive Theory In Multilateral Decentralized Markets," *Journal of Political Economy* 112(5), 1131-1156.
- Loken, E. and A. Gelman. (2017) "Measurement Error and the Replication Crisis," *Science* 355(6325): 584-585.
- Maniadis, Z., F. Tufano, and J. List. (2014) "One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects," *American Economic Review* 104(1), 277-90.
- Nuzzo, S. and A. Morone (2017) "Asset Markets In The Lab: A Literature Review," *Journal of Behavioral and Experimental Finance* 13, 42-50.
- Noussair, C. and S. Tucker (2013) "A Collection of Surveys on Market Experiments," Wiley Blackwell.
- Page, L. and C. Siemroth (2017) "An Experimental Analysis of Information Acquisition in Prediction Markets," *Games and Economic Behavior* 101, 354-378.
- Page, L. and C. Siemroth (2018) "How Much Information is Incorporated in Financial Asset Prices? Experimental Evidence," Working Paper.
- Pennock, D. S. Lawrence, C.L. Lee, and F. A. Nielsen (2001) "The Real Power of Artificial Markets," *Science* 291, 987-988.
- Radner, R. (1982) "Equilibrium under Uncertainty," in Handbook of Mathematical Economics II, North-Holland Publishing Company.
- Rud, O., J. P. Rabanal, and M. Sharifova (2019) "An Experiment of the Efficiency of Bilateral Exchange Under Incomplete Markets", *Games and Economic Behavior* 114, 253-267.
- Shiller, R. (2015) "Irrational Exuberance," 3rd edition, Princeton University Press.
- Shleifer, A. (2000) "Clarendon Lectures: Inefficient Markets," Oxford University Press.
- Smith V. (1962) "An Experimental Study of Competitive Market Behavior," *The Journal of Political Economy* 70, 111-137.
- Tetlock, P. (2004) "How Efficient are Information Markets? Evidence from an Online Exchange," Harvard University, Working Paper.
- Thaler, R. (2005) "Advances in Behavioral Finance," Volume II (Roundtable Series in Behavioral Economics), Princeton: Princeton University Press.
- Thaler, R. (2015) "Misbehaving: The Making of Behavioral Economics," New York: W. W. Norton & Company.
- Veiga, H. and M. Vorsatz (2010) "Information Aggregation in Experimental Asset Markets in The Presence of a Manipulator," *Experimental Economics* 13(4), 379-398.

9. Appendices

Appendix A. Prediction Tables

Price predictions for Series B and C Markets are shown in Table A1.

Table A1. Price predictions for each model (Due to P&S).

Market	Model	True State		
		X	Y	Z
4X	RE	230	0	0
	PI	146	146	101
	MM	0	0	0
4Y	RE	0	160	0
	PI	58	58	49
	MM	0	0	0
4Z	RE	0	0	300
	PI	169	208	208
	MM	0	0	0
5X	RE	460	0	0
	PI	230	230	230
	MM	0	0	0
5Y	RE	0	320	0
	PI	160	160	160
	MM	0	0	0
5Z	RE	0	0	600
	PI	300	300	300
	MM	0	0	0
7	RE	50	240	590
	PI	270	365	365
	MM	50	240	240
8	RE	125	375	450
	PI	325	450	450
	MM	125	375	375
9	RE	50	240	490
	PI	210	317	317
	MM	50	240	240

Allocation predictions for Series B and C Markets are shown in Table A2.

Table A2. Allocation predictions for each model (Due to P&S).
Type (I, II and III) and clue of traders who should hold the certificates according to a given model.

Market	Model	True State		
		X	Y	Z
4X	RE	II	–	–
	PI	II (Not Z)	II (Not Z)	II (Not Y)
	MM	-----No predictions-----		
4Y	RE	–	III	–
	PI	III (Not Z)	III (Not Z)	III (Not X)
	MM	-----No predictions-----		
4Z	RE	–	–	I
	PI	I (Not Y)	I (Not X)	I (Not X)
	MM	-----No predictions-----		

5X	RE	II	–	–
	PI	II	II (Not Z)	II (Not Y)
	MM	-----No predictions-----		
5Y	RE	–	III	–
	PI	III (Not Z)	III	III (Not X)
	MM	-----No predictions-----		
5Z	RE	–	–	I
	PI	I (Not Y)	I (Not X)	I
	MM	-----No predictions-----		

7	RE		All (no trade)	
	PI	Not Y	Not Y	Not Y
	MM	All (no trade)	Not X	Not X
8	RE		All (no trade)	
	PI	Not Y	Not X	Not X
	MM	All (no trade)	Not X	Not X
9	RE		All (no trade)	
	PI	Not Y	Not Y	Not Y
	MM	All (no trade)	Not X	Not X

Price predictions for Series D Markets are shown in Table A3.

Table A3. Price predictions for each model.

Market	Model	True State		
		X	Y	Z
12X	RE	125	0	0
	PI	62.5	62.5	62.5
	MM	0	0	0
12Y	RE	0	375	0
	PI	187.5	187.5	187.5
	MM	0	0	0
12Z	RE	0	0	525
	PI	262.5	262.5	262.5
	MM	0	0	0
13X	RE	50	0	0
	PI	31.8	21.9	31.8
	MM	0	0	0
13Y	RE	0	240	0
	PI	135	166.2	166.2
	MM	0	0	0
13Z	RE	0	0	490
	PI	178.2	150.8	178.2
	MM	0	0	0

Allocation predictions for Series D Markets are shown in Table A4.

Table A4. Allocation predictions for each model.
Clue of traders who should hold the certificates according to a given model.

Market	Model	True State		
		X	Y	Z
12X	RE	-----No predictions-----		
	PI	All (no trade)	Not Z	Not Y
	MM	-----No predictions-----		
12Y	RE	-----No predictions-----		
	PI	Not Z	All (no trade)	Not X
	MM	-----No predictions-----		
12Z	RE	-----No predictions-----		
	PI	Not Y	Not X	All (no trade)
	MM	-----No predictions-----		
13X	RE	-----No predictions-----		
	PI	Not Y	Not Z	Not Y
	MM	-----No predictions-----		
13Y	RE	-----No predictions-----		
	PI	Not Z	Not X	Not X
	MM	-----No predictions-----		
13Z	RE	-----No predictions-----		
	PI	Not Y	Not X	Not Y
	MM	-----No predictions-----		

Appendix B. Details for Conduct of Experimental Sessions and Complete Subject Instructions

Description of Stages of a Session

Stage 1. Subjects were trained with the mechanism (a bingo cage) used to draw the state of nature. In Plott and Sunder (1982, p. 693), it is noted that 10-20 draws were completed until most subjects predicted the most likely outcome. In our replication, there were always exactly 10 draws for this training.

Stage 2. Subjects were trained with the mechanism to provide signals (referred to as clues) about the state of nature. P&S note (p. 1118) that this training lasted between 8 and 10 trials. In our replication, there were always exactly 10 draws for this training. Subjects were given a *Clue Sheet* with 30 rows. The first 10 rows were used for training, and then their clue for the first market period was on row 11. This ensured that the number of possible periods on their *Profit Sheets* matched the number of remaining rows in the *Clue Sheet*, post-training.

Stage 3. Subjects were given an explanation of the market rules. Because we wanted to use the same sequence of state realizations as in P&S and did not want to deceive the subjects, rather than continuing to randomly draw the state with a bingo cage as in Stage 1, our subjects were informed that “The outcomes were determined before the experiment began by drawing a ball from a bingo cage in a manner similar to what was described earlier.” The subjects were not informed that the draws occurred thirty years before the session. We also went through an example sequence of bids, asks and acceptances. The example varied based on whether one or multiple assets were available for trade. As in P&S, subjects could hold short positions during trading, but incurred a penalty equal to 300 francs plus the highest transaction price in that period per unit short at the end of a trading period. This contrasts with Plott and Sunder (1982, p. 696) where short positions were not allowed, and hence this necessitated a modification to the instructions.

Stage 4. The market was conducted for several periods, referred to as years in the experiment. Each period lasted for seven minutes, and all bids, asks and contracts were left visible on the board until the period ended. We conducted the sessions in rooms with ample whiteboard space, and we never had to erase information mid-period. The *Information and Record Sheets* were collected at the end of each period, and new *Information and Record Sheets* were handed out before the start of each period - subjects kept their *Profit Sheet* throughout Stage 4. In our replication, subjects were allowed to use basic four-function calculators to update their *Information and Record Sheets* and *Profit Sheet*. Bids, asks and contracts from previous periods were erased before the next period began. Subjects did not know how many periods of trading there would be, as the *Profit Sheet* always listed 20 periods.

Stage 5. The subjects received their payment in private and were dismissed from the experiment.

Notes: In Market 4 where the series type changes mid-session, the subjects repeated Stage 3 after period 9. In Markets 5 and 8, which involve only experienced subjects and were conducted as brand-new sessions, we repeated all five stages.

Steps to Run Hand Run Information Aggregation Markets

Prior to Experiment

1. Set up Bingo cage with correct number of balls and have a coin available.
2. Have market program set up on podium computer (not projected). Have document camera for Market Operation Example.
3. Place a numbered (1 through 12) notecard at each position asking for subject name and student id.
4. Place a) calculator, b) pen, c) Trader Sign, and d) Subject Name and ID Card at each position.
5. Place a) Instruction Set 1, b) Response Form for random draw training, and c) Clue Sheets.
Note: a) and c) are market specific.

Subject Sign in

6. Subjects consent
7. Subjects fill out Subject Name and ID card and RA signs them into the market program

Experiment

8. Auctioneer reads Script 0
9. Auctioneer reads Script 1
10. Conduct 10 Drawings with replacement
11. Collect Response Forms
12. Auctioneer reads Script 2
13. Auctioneer conducts 10 practice rounds with clues.
 - a) draw a ball
 - b) flip coin
 - c) call out row (=practice number) and column (based on draw and flip)
 - d) put ball back in bingo cage
14. Handout a) Instruction Set 3, b) Profit Sheet, and c) Year 1 Information and Record Sheet. Note: b) is market specific while a) and c) are market and trader specific.
15. Auctioneer reads Instruction Set 3. “___” is read as blank because this is private information for each trader.
16. Auctioneer goes through the Market Operation Example.
Order of call should be Trader #, Action, (Market), Price. (Market) is only for contingent claims.
17. Conduct Market Year
 - a) Handout Year Y record and Information Sheet (already done for Year 1). Note: This is market and trader specific
 - b) Announce Clue row and column. Row is 10+year. Column indicated by “pickme.”
 - c) Start 7 minute market year.
 - d) Repeat a-c until last Year completed. Number of Years differs by Market. For Market 4, return to Step 14 after Year 9.
18. Collect all papers. Calculate total payments = Trading Profit + Prediction Payment from 12 + Show up Payment.
19. Pay and dismiss subjects.

Set 0 Script

This is an experiment in the economics of market decision making. Various research foundations have provided funds for this research. The instructions are simple, and if you follow them carefully and make good decisions, you might earn a considerable amount of money which will be paid to you in cash. In this experiment we are going to simulate a market in which you will buy and sell certificates in a sequence of market years.

This experiment has four steps:

- (1) training with the mechanism used to draw states of nature,
- (2) training with the mechanism used to provide clues about the state of nature,
- (3) explanation of procedures and rules of the market,
- (4) conduct of markets for several years.

Set 1 Script

Instruction Set 1 [Market 4]

Each year we draw a ball from a bingo cage containing twenty balls numbered one through twenty. If the ball drawn is numbered one through seven, outcome of the draw is called x ; if a ball numbered eight through eleven is drawn, the outcome is called y ; if a ball numbered twelve through twenty is drawn, the outcome is called z .

You have to predict the outcome of each draw before it is announced. If your prediction is correct, you win \$0.50; if wrong, you lose \$0.20. Before the first draw is made, record your prediction by circling either x or y or z in the first row of the enclosed sheet. After you have encircled one letter, the outcome will be announced and you should record the announced outcome in the blank space on the same row of the table. If your prediction is correct, circle the amount shown in the Win column, otherwise circle the amount shown in the Lose column.

Once you have recorded your prediction you must not make a change; any erasure will invalidate your prediction. At the end, add up your total winnings and losses and record the difference (net winnings or losses) at the bottom right corner of the sheet.

Instruction Set 1 [Markets 9 & 13]

Each year we draw a ball from a bingo cage containing twenty balls numbered one through twenty. If the ball drawn is numbered one through seven, outcome of the draw is called x; if a ball numbered eight through sixteen is drawn, the outcome is called y; if a ball numbered seventeen through twenty is drawn, the outcome is called z.

You have to predict the outcome of each draw before it is announced. If your prediction is correct, you win \$0.50; if wrong, you lose \$0.20. Before the first draw is made, record your prediction by circling either x or y or z in the first row of the enclosed sheet. After you have encircled one letter, the outcome will be announced and you should record the announced outcome in the blank space on the same row of the table. If your prediction is correct, circle the amount shown in the Win column, otherwise circle the amount shown in the Lose column.

Once you have recorded your prediction you must not make a change; any erasure will invalidate your prediction. At the end, add up your total winnings and losses and record the difference (net winnings or losses) at the bottom right corner of the sheet.

Instruction Set 1 [Markets 5, 7, 8, & 12]

Each year we draw a ball from a bingo cage containing thirty balls numbered one through thirty. If the ball drawn is numbered one through ten, outcome of the draw is called x; if a ball numbered eleven through twenty is drawn, the outcome is called y; if a ball numbered twenty-one through thirty is drawn, the outcome is called z.

You have to predict the outcome of each draw before it is announced. If your prediction is correct, you win \$0.50; if wrong, you lose \$0.20. Before the first draw is made, record your prediction by circling either x or y or z in the first row of the enclosed sheet. After you have encircled one letter, the outcome will be announced and you should record the announced outcome in the blank space on the same row of the table. If your prediction is correct, circle the amount shown in the Win column, otherwise circle the amount shown in the Lose column.

Once you have recorded your prediction you must not make a change; any erasure will invalidate your prediction. At the end, add up your total winnings and losses and record the difference (net winnings or losses) at the bottom right corner of the sheet.

States of Nature Training Sheet

Subject No. _____

Number	Circle One Decision	Outcome x or y or z	Win (\$)	Lose (\$)
1	x y z	_____	0.50	-0.20
2	x y z	_____	0.50	-0.20
3	x y z	_____	0.50	-0.20
4	x y z	_____	0.50	-0.20
5	x y z	_____	0.50	-0.20
6	x y z	_____	0.50	-0.20
7	x y z	_____	0.50	-0.20
8	x y z	_____	0.50	-0.20
9	x y z	_____	0.50	-0.20
10	x y z	_____	0.50	-0.20

Total winnings _____

Total losses _____

Net winnings/losses _____

Set 2 Script

You have each received a clue sheet. Do not reveal this sheet to anyone.

You will use this clue sheet to receive a clue about what the outcome is in a particular year. Therefore, we will now go through some examples of how to use the clue sheet.

Each year the experimenter will draw a bingo ball from the cage to determine if the outcome is x or y or z . The experimenter will then call out a row and column. You will use this row and column to determine your clue. The other investors will use their own sheets to determine their clues.

If the outcome is x then there are two possible clues: not y and not z

If the outcome is y then there are two possible clues: not x and not z

If the outcome is z then there are two possible clues: not x and not y

Everyone's sheet is different. Which of the two possible clues that you receive is determined randomly. The clues that other investors receive are also determined randomly. But the clues are distributed in such a way that exactly half of all investors receive one of the two possible clues and the other half receive the other possible clue.

We will now practice with the clue sheet.

Clue Sheet [Each clue sheet was trader and Market Specific]

Clue Sheet

Trader #: _____

Columns						
	1	2	3	4	5	6
1	not z	not x	not y	not z	not y	not x
2	not z	not x	not x	not y	not z	not y
3	not x	not y	not x	not z	not y	not z
4	not z	not y	not y	not x	not z	not x
5	not z	not x	not z	not y	not y	not x
6	not x	not x	not y	not y	not z	not z
7	not x	not x	not y	not z	not z	not y
8	not y	not z	not y	not x	not x	not z
9	not y	not z	not x	not x	not z	not y
10	not z	not y	not z	not y	not x	not x
11	not x	not z	not z	not y	not x	not y
12	not z	not z	not y	not y	not x	not x
13	not y	not x	not x	not z	not y	not z
14	not z	not z	not x	not y	not y	not x
15	not x	not y	not y	not z	not z	not x
16	not y	not x	not x	not z	not z	not y
17	not z	not z	not y	not y	not x	not x
18	not x	not y	not z	not x	not z	not y
19	not y	not z	not z	not x	not y	not x
20	not y	not y	not z	not z	not x	not x
21	not z	not y	not x	not z	not y	not x
22	not x	not z	not y	not z	not x	not y
23	not x	not x	not z	not y	not y	not z
24	not x	not x	not y	not z	not y	not z
25	not z	not y	not x	not z	not x	not y
26	not z	not y	not x	not z	not y	not x
27	not x	not x	not y	not z	not y	not z
28	not y	not x	not x	not z	not z	not y
29	not x	not y	not z	not z	not x	not y
30	not y	not y	not x	not z	not z	not x

Set 3 Script

Instructions Set 3 Script [Single Asset; State Probabilities for Markets 9 and 13]

General. – This is an experiment in the economics of market decision making. Various research foundations have provided funds for this research. The instructions are simple, and if you follow them carefully and make good decisions, you might earn a considerable amount of money which will be paid to you in cash.

In this experiment we are going to simulate a market in which you will buy and sell certificates in a sequence of market years. Attached to the instructions you will find a sheet, labeled information and record sheet, which helps determine the value to you of any decisions you might make. You are not to reveal this information to anyone. It is your own private information.

The type of currency used in this market is francs. All trading and earnings will be in terms of francs. Each franc is worth ____ to you. Do not reveal this number to anyone. At the end of the experiment your francs will be converted to dollars at this rate, and you will be paid in dollars. Notice that the more francs you earn the more dollars you earn.

Specific Instructions

Your profits come from two sources—from collecting certificate earnings on all certificates you hold at the end of the year and from buying and selling certificates. During each market year you are free to purchase or sell as many certificates as you wish, provided you follow the rules below. For each certificate you hold at the end of the year you will be given one of the three numbers of francs listed on row 19 of your information and record sheet. Note that earnings may be different for different investors. The method by which one of the three numbers is selected each year is explained later in these instructions. Compute your total certificate earnings for a period by multiplying the earnings per certificate by the number of certificates held. That is, $(\text{number of certificates held}) \times (\text{earnings per certificate}) = \text{total certificate earnings}$. Suppose, for example, that you hold five certificates at the end of year 1. If for that period your earnings are ____ francs per certificate (i.e., the number selected from row 19 is ____) then your total certificate earnings in the year would be $5 \times \text{____} = \text{____}$ francs. This number should be recorded on row 19 at the end of the year.

Sales from your certificate holdings increase your francs on hand by the amount of the sale price. Similarly, purchases reduce your francs on hand by the amount of the purchase price. Thus you can gain or lose money on the purchase and resale of certificates. At the end of each year all your holdings are automatically sold to the experimenter at a price of 0. At the beginning of each year you are provided with an initial holding of certificates. This is recorded on row 0 of the year's information and record sheet. You may sell these if you wish or you may

hold them. If you hold a certificate, then you receive "earnings per certificate" at the end of the year.

In addition, at the beginning of each year you are provided with an initial amount of francs on hand. This is also recorded on row 0 of each year's information and record sheet. You may keep this if you wish or you may use it to purchase certificates.

Thus at the beginning of each year you are endowed with holdings of certificates and francs on hand. You are free to buy and sell certificates as you wish according to the rules below. Your francs on hand at the end of a year are determined by your initial amount of francs on hand, earnings on certificate holdings at the end of the year, and by gains and losses from purchases and sales of certificates. All francs on hand at the end of a year in excess of _____ francs are yours to keep. These are your profits for the year.

Information about Dividends

Whether the dividend you receive from the certificates you hold is the x-dividend shown on row 19 or the y-dividend on row 19 or the z-dividend on row 19 depends on the draw of an outcome from a bingo cage. If the outcome is x the dividend is the x-dividend. If the outcome is y the dividend is the y-dividend. If the outcome is z the dividend is the z-dividend. The outcomes were determined before the experiment began by drawing a ball from a bingo cage in a manner similar to what was described earlier.

As a reminder, each year there is a 35% chance that the outcome is x; a 45% chance that the outcome is y; and a 20% chance that the outcome is z.

At the beginning of each year, before trading starts, each investor will receive a clue using the clue sheets you have already received.

Trading and Recording Rules

- (1) All transactions are for one certificate at a time. After each of your sales or purchases you must record the TRANSACTION PRICE in the appropriate column depending on the nature of the transaction. The first transaction is recorded on row (1), and succeeding transactions are recorded on subsequent rows.
- (2) After each transaction you must calculate and record your new holdings of certificates and your new francs on hand. Your holdings of certificates may go below zero. Your francs on hand may never go below zero.
- (3) At the end of the year record your total certificate earnings in the last column of row 19. If you have negative certificate holdings, you receive no dividends, and you must pay a penalty of 300 francs plus the highest transaction price during the period, for each

certificate you are short. Compute your end of period totals on row 20 by listing certificate holdings and adding total certificate earnings to your francs on hand.

- (4) At the end of the year, subtract from your francs on hand the amount listed in row 21 and enter this new amount on row 22. This is your profit for the market year and is yours to keep. At the end of each market year, record this number on your profit sheet.
- (5) At the end of the experiment add up your total profit on your profit sheet and enter this sum on row 22 of your profit sheet. To convert this number into dollars, multiply by the number on row 23 and record the product on row 24. The experimenter will pay you this amount of money.

Market organization. – The market for these certificates is organized as follows. The market will be conducted in a series of years. Each period lasts for seven minutes. Anyone wishing to purchase a certificate is free to raise his or her hand and make a verbal bid to buy one certificate at a specified price, and anyone with certificates to sell is free to accept or not accept the bid. Likewise, anyone wishing to sell a certificate is free to raise his or her hand and make a verbal offer to sell one certificate at a specified price. If a bid or offer is accepted, a binding contract has been closed for a single certificate, and the contracting parties will record the transaction on their information and record sheets. Any ties in bids or acceptance will be resolved by random choice. Except for the bids and their acceptance, you are not to speak to any other subject. There are likely to be many bids that are not accepted, but you are free to keep trying. You are free to make as much profit as you can.

Instructions Set 3 Script [Contingent Claims; State Probabilities for Markets 9 and 13]

General. – This is an experiment in the economics of market decision making. Various research foundations have provided funds for this research. The instructions are simple, and if you follow them carefully and make good decisions, you might earn a considerable amount of money which will be paid to you in cash.

In this experiment we are going to simulate a market in which you will buy and sell certificates in a sequence of market years. Attached to the instructions you will find a sheet, labeled information and record sheet, which helps determine the value to you of any decisions you might make. You are not to reveal this information to anyone. It is your own private information.

The type of currency used in this market is francs. All trading and earnings will be in terms of francs. Each franc is worth ____ to you. Do not reveal this number to anyone. At the end of the experiment your francs will be converted to dollars at this rate, and you will be paid in dollars. Notice that the more francs you earn the more dollars you earn.

Specific Instructions

Your profits come from two sources—from collecting certificate earnings on all certificates you hold at the end of the year and from buying and selling certificates. During each market year you are free to purchase or sell as many certificates as you wish, provided you follow the rules below. There are three types of certificate: x-certificate, y-certificate, and z-certificate. For each certificate of a given type that you hold at the end of the year you will be given either the number of francs listed on row 19 of your information and record sheet for that certificate type or 0. Note that earnings may be different for different investors. The method by which one of the two numbers for each certificate type is selected each year is explained later in these instructions. Compute your total certificate earnings for a period by multiplying the earnings per certificate of a given type by the number of certificates of that type held and then summing these three amounts. That is, (number of x-certificates held) x (earnings per x-certificate) + (number of y-certificates held) x (earnings per y-certificate) + (number of z-certificates held) x (earnings per z-certificate) = total certificate earnings. Suppose, for example, that you hold five x-certificates at the end of year 1. If for that period your earnings are ____ francs per x-certificate (i.e., the number for x-certificates from row 19 is ____) then your total x-certificate earnings in the year would be $5 \times \text{____} = \text{____}$ francs. Suppose you hold 4 y-certificates and 3 z-certificates at the end of year 1. If for that period your earnings are 0 francs per certificate for y-certificates and z-certificates then your total y-certificate earnings in the year would be $4 \times 0 = 0$ francs and your total z-certificate earnings in the year would be $3 \times 0 = 0$ francs. Your total certificate earnings would be $\text{____} + 0 + 0 = \text{____}$ francs. This number should be recorded on row 19 at the end of the year.

Sales from your certificate holdings increase your francs on hand by the amount of the sale price. Similarly, purchases reduce your francs on hand by the amount of the purchase price. Thus you can gain or lose money on the purchase and resale of certificates. At the end of each year all your holdings are automatically sold to the experimenter at a price of 0. At the beginning of each year you are provided with an initial holding of certificates. This is recorded on row 0 of the year's information and record sheet. You may sell these if you wish or you may hold them. If you hold a certificate, then you receive "earnings per certificate" at the end of the year.

In addition, at the beginning of each year you are provided with an initial amount of francs on hand. This is also recorded on row 0 of each year's information and record sheet. You may keep this if you wish or you may use it to purchase certificates.

Thus at the beginning of each year you are endowed with holdings of certificates and francs on hand. You are free to buy and sell certificates as you wish according to the rules below. Your francs on hand at the end of a year are determined by your initial amount of francs on hand, earnings on certificate holdings at the end of the year, and by gains and losses from purchases and sales of certificates. All francs on hand at the end of a year in excess of _____ francs are yours to keep. These are your profits for the year.

Information about Dividends

Whether the dividend you receive from the certificates of a given type you hold is the amount shown on row 19 or 0 depends on the draw of an outcome from a bingo cage. If the outcome is x then x -certificates receive the x -dividend shown on row 19 while y -certificates and z -certificates each earn 0. If the outcome is y then y -certificates receive the y -dividend shown on row 19 while x -certificates and z -certificates each earn 0. If the outcome is z then z -certificates receive the z -dividend shown on row 19 while x -certificates and y -certificates each earn 0. The outcomes were determined before the experiment began by drawing a ball from a bingo cage in a manner similar to what was described earlier.

As a reminder, each year there is a 35% chance that the outcome is x ; a 45% chance that the outcome is y ; and a 20% chance that the outcome is z .

At the beginning of each year, before trading starts, each investor will receive a clue using the clue sheets you have already received.

Trading and Recording Rules

- (6) All transactions are for one certificate of a given type at a time. After each of your sales or purchases you must record the TRANSACTION PRICE in the appropriate column

depending on the nature of the transaction. The first transaction is recorded on row (1), and succeeding transactions are recorded on subsequent rows.

- (7) After each transaction you must calculate and record your new holdings of certificates and your new francs on hand. Your holdings of certificates of any type may go below zero. Your francs on hand may never go below zero.
- (8) At the end of the year record your total certificate earnings in the last column of row 19. If you have negative certificate holdings of a type, you receive no dividends for that type, and you must pay a penalty of 300 francs plus the highest transaction price for that type during the period, for each certificate you are short. Compute your end of period totals on row 20 by listing certificate holdings and adding total certificate earnings to your francs on hand.
- (9) At the end of the year, subtract from your francs on hand the amount listed in row 21 and enter this new amount on row 22. This is your profit for the market year and is yours to keep. At the end of each market year, record this number on your profit sheet.
- (10) At the end of the experiment add up your total profit on your profit sheet and enter this sum on row 22 of your profit sheet. To convert this number into dollars, multiply by the number on row 23 and record the product on row 24. The experimenter will pay you this amount of money.

Market organization. – The market for these certificates is organized as follows. The market will be conducted in a series of years. Each period lasts for seven minutes. Anyone wishing to purchase a certificate of a given type is free to raise his or her hand and make a verbal bid to buy one certificate of a given type at a specified price, and anyone with certificates of the given type to sell is free to accept or not accept the bid. Likewise, anyone wishing to sell a certificate of a given type is free to raise his or her hand and make a verbal offer to sell one certificate of a given type at a specified price. If a bid or offer is accepted, a binding contract has been closed for a single certificate, and the contracting parties will record the transaction on their information and record sheets. Any ties in bids or acceptance will be resolved by random choice. Except for the bids and their acceptance, you are not to speak to any other subject. There are likely to be many bids that are not accepted, but you are free to keep trying. You are free to make as much profit as you can.

Name _____

PROFIT SHEET

Row	Market Year	Profit
1	1	
2	2	
3	3	
4	4	
5	5	
6	6	
7	7	
8	8	
9	9	
10	10	
11	11	
12	12	
13	13	
14	14	
15	15	
16	16	
17	17	
18	18	
19	19	
20	20	
21		
22	Total profit (in francs)	
23	Dollars per franc	0.003 ²⁴
24	Total dollars profit	

²⁴ This exchange rate varies across markets.

Trader No. _____

INFORMATION AND RECORD SHEET [Single Asset]

YEAR _____

	Transaction Number	Transaction Price		Certificates on Hand	Francs on Hand
		Sale	Purchase		
Beginning of the Year Holdings	0	////////////////////			
	1				
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				
	10				
	11				
	12				
	13				
	14				
	15				
	16				
	17				
	18				
x-Dividend _____	19	Total Certificate Earnings Dividend Rate on Hand at the End of the Year			
y-Dividend _____					
z-Dividend _____					
	20	Total Francs on Hand at the End of the Year			
	21	Less: Fixed Costs			
	22	End of Year Net Profit			

Transfer this amount
to your Profit Sheet

Trader No. _____

INFORMATION AND RECORD SHEET [Contingent Claims]

YEAR _____

Beginning of the
Year Holdings

Transaction Number	Transaction Price		Certificates on Hand			Francs on Hand
	Sale	Purchase	x	y	z	
0	////////////////////					
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19	Total Certificate Earnings Dividend Rate on Hand at the End of the Year					
20	Total Francs on Hand at the End of the Year					
21	Less: Fixed Costs					
22	End of Year Net Profit					

x-Dividend _____
y-Dividend _____
z-Dividend _____

Transfer this amount
to your Profit Sheet



Demonstration of Market [Single Asset]

We will now go through a demonstration of how the market will work.

The auctioneer will announce the market is open.

Suppose that Trader 13 wants to offer to buy a certificate at a price of 25. Trader 13 would raise her paddle and the auctioneer would call on her saying something like "Trader 13." Trader 13 would then respond with "Bid of 25." The auctioneer will repeat the order: "Trader 13 Bids 25." This will be recorded on the board. [The order is marked on the board.]

Now suppose that Trader 14 wants to offer to sell a certificate at a price of 500. Trader 14 should raise his paddle. When the auctioneer calls on Trader 14, Trader 14 should announce "Ask 500." The auctioneer will repeat "Trader 14 Asks 500." This will then be recorded on the board. [The order is marked on the board.]

This process will continue as long as someone wants to improve on the currently standing bid or ask or accept the standing bid or ask. For example, Trader 13 may raise her paddle and after being called state "Bid of 100." The auctioneer will repeat "Trader 13 Bids 100" and this will be marked on the board. [The order is marked on the board.] Notice that you can improve upon your own bid or ask.

If Trader 15 is willing to sell at 100, the current standing bid, Trader 15 can raise her paddle. When called upon Trader 15 will state "Accept Bid 100." The auctioneer will repeat that "Trader 15 Accepts Bid of 100." At this point Trader 13 and 15 have traded a single certificate and this will be reflected on the board by circling the accepted bid. [Circle the accepted bid on the board and write Trader 15 on the line as the seller.]

Trader 13 will record on her Information and Record sheet that she purchased a certificate by writing the price in the Purchase column in the row corresponding to Transaction 1. She will then update her Certificates on Hand and Francs on Hand. Similarly, Trader 15 will record on her Information and Record sheet that she sold a certificate by writing the price in the Sale column in the row corresponding to Transaction 1. She will then update her Certificates on Hand and Francs on Hand.

A line will be drawn across the market record on the board indicating that all of the previous bids and asks are no longer available. The auctioneer will then invite new bids or asks.

Trader 13 may state an Ask of 600.

Trader 14 may state an Ask of 542.

Trader 16 may state a Bid of 33.

Trader 17 may state a Bid of 35.

Trader 16 may state a bid of 40.

Trader 17 may state a Bid of 60.

Trader 14 may state an Ask 500.

Trader 15 may Accept the Bid of 60.

At this point, the bid of 60 is circled and Traders 17 would record the price of 60 under Purchase in the row for Transaction 1. Trader 17 would then update his Certificates on Hand and Francs on Hand. Notice that for Trader 15 this would be her second transaction, so she would record the Sale price on the row for Transaction 2 and then update the remainder of the row.

A line will be drawn across the market record on the board indicating that all of the previous bids and asks are no longer available and the auctioneer will again invite new bids or asks. This process continues until 7 minutes have passed at which point no more bids, asks, or acceptance can be made.

Suppose there are no more transactions involving Trader 15 in the market year. If the state is Y, given the fictitious values in this example, Trader 15 would complete their Information and Record Sheet as follows:

Total Certificate Earnings in Row 19 would be: $8 * 100 = 800$

Total Francs on Hand in Row 20 would be: 20,160

End of Year Net Profit in Row 22 would be: $800 + 20,160 - 20,000 = 960$.

Demonstration of Market [Contingent Claims]

We will now go through a demonstration of how the market will work.

The auctioneer will announce the market is open.

Suppose that Trader 13 wants to offer to buy an X certificate at a price of 25. Trader 13 would raise her paddle and the auctioneer would call on her saying something like "Trader 13." Trader 13 would then respond with "Bid on X of 25." The auctioneer will repeat the order: "Trader 13 Bid on X of 25." This will be recorded on the board. [The order is marked on the board.]

Now suppose that Trader 14 wants to offer to sell an X certificate at a price of 500. Trader 14 should raise his paddle. When the auctioneer calls on Trader 14, Trader 14 should announce "Ask for X of 500." The auctioneer will repeat "Trader 14 Ask for X of 500." This will then be recorded on the board. [The order is marked on the board.]

This process will continue as long as someone wants to improve on the currently standing bid or ask or accept the standing bid or ask. For example, Trader 13 may raise her paddle and after being called state "Bid on X of 100." The auctioneer will repeat "Trader 13 Bid on X of 100" and this will be marked on the board. [The order is marked on the board.] Notice that you can improve upon your own bid or ask.

If Trader 15 is willing to sell X at 100, the current standing bid, Trader 15 can raise her paddle. When called upon Trader 15 will state "Accept Bid on X of 100." The auctioneer will repeat that "Trader 15 Accepts Bid on X of 100." At this point Trader 13 and 15 have traded a single X certificate and this will be reflected on the board by circling the accepted bid. [Circle the accepted bid on the board and write Trader 15 on the line as the seller.]

Trader 13 will record on her Information and Record sheet that she purchased an X certificate by writing the price in the Purchase column in the row corresponding to Transaction 1. She will then update her Certificates on Hand and Francs on Hand. Similarly, Trader 15 will record on her Information and Record sheet that she sold an X certificate by writing the price in the Sale column in the row corresponding to Transaction 1. She will then update her Certificates on Hand and Francs on Hand.

A line will be drawn across the X market record on the board indicating that all of the previous bids and asks for X certificates are no longer available. The auctioneer will then invite new bids or asks.

Trader 13 may state an Ask for X of 600.

Trader 14 may state an Ask for Y of 542.

Trader 17 may state a Bid on Y of 35

Trader 16 may state a Bid on X of 33.

Trader 16 may state a bid on X of 40.

Trader 17 may state a Bid on Y of 60.

Trader 14 may state an Ask on X for 500.

Trader 15 may Accept the Bid on Y of 60.

At this point, the bid of 60 for Y is circled and Trader 17 would record the price of 60 under Purchase in the row for Transaction 1. Trader 17 would then update all of his Certificates on Hand and Francs on Hand. Notice that for Trader 15 this would be her second transaction, so she would record the Sale price on the row for Transaction 2 and then update the remainder of the row.

A line will be drawn across the market record for Y on the board indicating that all of the previous bids and asks for Y are no longer available and the auctioneer will again invite new bids or asks. Notice that the bids and asks for X are still available. This process continues until 7 minutes have passed at which point no more bids, asks, or acceptance can be made.

Suppose there are no more transactions involving Trader 15 in the market year. If the state is Y, given the fictitious values in this example, Trader 15 would complete their Information and Record Sheet as follows:

Total Certificate Earnings in Row 19 would be: $9 * 0 + 9 * 100 + 10 * 0 = 900$

Total Francs on Hand in Row 20 would be: 20,160

End of Year Net Profit in Row 22 would be: $900 + 20,160 - 20,000 = 1,060$.

Appendix C.

Additional Evidence on Series B and C and Graphs of Transaction Prices by Sessions for Markets 4, 5, 7, and 8

Allocations

As in P&S, we proceed by assessing whether the allocation of shares at the end of each period (corresponding to the last occurrence of each state of nature) is consistent with the allocation predicted by each model (see Appendix A or P&S), and we compare the fit across models. The allocation statistic is determined via the formula:

$$A = \left(\sum_{c \in C} \sum_{i \in I_c} n_{i,c} - \bar{n}_c \right) / \left(\sum_{c \in C} \sum_{i \in I_c} n_{i,c}^{(m)} - \bar{n}_c \right) \times 100$$

where $n_{i,c}$ represents the number of shares of certificates c held by investor i at the end of the market, I_c denotes the set of traders predicted to hold the asset, C denotes the set of certificates in the market, $n_{i,c}^{(m)}$ corresponds to the number of shares predicted to be held by investor i according to model m , and \bar{n} is set to the average number of shares initially held by all investors in I_c . This ratio is then averaged over the three periods in which each state was realized for the last time. The values for each replicate of each session are shown in Table C1 and can be compared to Table IV in P&S. The RE model provides no prediction for Series C because all traders end up learning the true asset value and thus are only willing to trade at this exact value. The MM model provides no allocation predictions in Series B because all traders value the asset at 0.

In line with our previous findings on price predictions, the RE model does not outperform either PI or MM. These findings contradict Conclusion 3 in P&S (p. 1106) according to which “In all series, allocations aggregated over the final occurrence of each state are more accurately predicted by the RE model than either the PI or MM.” However, when looking at Series B only we do find that the RE model outperforms PI which is the only alternative model for which we have a prediction. This supports Finding 2 in the main text.

Table C1. Comparison of Actual Allocations with the Allocations Predicted by Three Models
 Criterion: Percent of Predicted Flow of Securities that Actually Occurred at the End of Each Market

Market Experiments			Models		
Series	Market	Replicate	PI	RE	MM
B	4	1	4.44	12.50	No Prediction
		2	3.89	12.50	No Prediction
		3	3.89	22.92	No Prediction
		4	9.45	12.50	No Prediction
	5	1	14.88	45.83	No Prediction
		2	16.67	50.00	No Prediction
		3	20.83	70.83	No Prediction
		4	14.83	52.08	No Prediction
C	7	1	31.94	No Prediction	41.67
		2	27.78	No Prediction	25.00
		3	9.21	No Prediction	27.08
		4	38.89	No Prediction	47.92
	8	1	5.56	No Prediction	45.84
		2	38.89	No Prediction	37.50
		3	52.78	No Prediction	29.17
		4	30.56	No Prediction	49.99
	9	1	40.28	No Prediction	52.09
		2	26.39	No Prediction	27.09
		3	23.61	No Prediction	25.00
		4	43.06	No Prediction	60.42
Wilcoxon Rank Sum Tests			↑	↑	↑
			↓	↓	
B-series Sessions			88	Undefined	
			0.035(RE)		
C-series Sessions			Undefined	Undefined	
All Sessions			265	135	
			0.203(RE)	0.487(MM)	

Reported metrics for each session are based on the last period in which the state was X, the last period in which the state was Y and the last period in which the state was Z, as applicable. Each statistical test compares RE to an alternative model. Each cell in the lower portion of the table contains the associated test statistic and level of significance. The model in parentheses is the one favored by the data based on the median. This table is comparable to Table IV in P&S. However, P&S report that a Wilcoxon Signed Rank Sum Test is used, but that test is not defined for any comparison between RE and another model except for a test between PI and RE in Series-B Markets. The test statistic for such a test would be 0 based on our data with a p-value of 0.012 with the data favoring RE.

Profits

Finally, we compare the models' predictions on the distribution of profits. In Table C2 (comparable to Table V in P&S), we display the weighted sum of the squared deviation between the average earnings of a trader of a given type and information and the average predicted earnings, given the trader's type and information, averaged over the three periods corresponding to the last occurrence of each state of nature.²⁵ In line with our previous findings on price and allocation predictions, we do not find that the RE model outperforms both alternatives; although, it significantly outperforms PI. These findings lead us to reject Conclusion 4 in P&S (p. 1108) according to which "In all series the RE model is a significantly better predictor of the distribution of profits than either the PI model or the MM model." On a positive note, when studied separately RE is shown to outperform the other two models in Series C markets. However, the profit prediction of the RE model for Series C markets is for every trader in the market to receive the same profit. Hence the RE model predicts that each person receives the average profit, which may drive some of the apparent success. A model that predicted no one ever placed a bid or ask in the market would have done equally well on this criterion for Series C. This evidence supports Finding 3 in the main text.

²⁵ Each model is silent with regards to which specific traders of a given type with a given piece of information should hold the assets. That is, if two traders of the same type, possessing the same private information, are predicted to purchase two assets being offered for sale, the models are silent as to whether one trader will buy both or each will buy one. Our approach thus removes profit deviations that are due to the allocation of assets within a given class of trader.

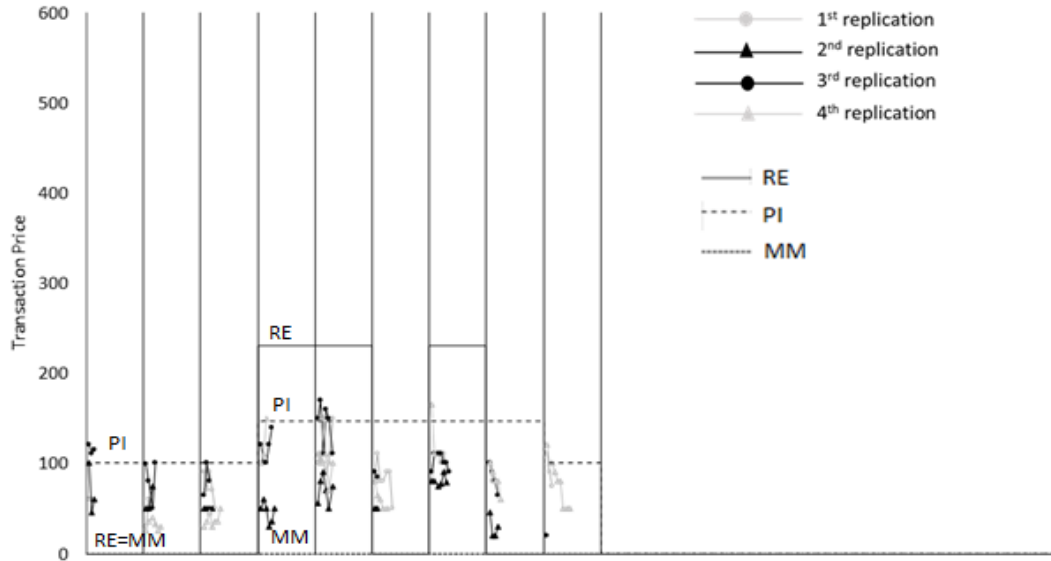
Table C2. Comparison of Actual Distribution of Profits to Distribution Predicted by Three Models

Criterion: Squared Sum of Deviations from the Mean Across Investors at the End of Each Market

Series	Market Experiments		Models		
	Market	Replicate	PI	RE	MM
B	4	1	8644.73	638.31	245.51
		2	8782.26	684.59	262.73
		3	8932.91	686.13	283.23
		4	8302.06	604.81	245.21
	5	1	24266.25	2721.85	1949.32
		2	25360.95	2474.61	2182.88
		3	24838.72	1912.72	1989.12
		4	25273.38	2858.71	2068.18
C	7	1	4281.22	49.62	3422.95
		2	4752.76	7.30	3787.30
		3	5473.51	36.55	4783.22
		4	3408.69	201.63	2894.96
	8	1	819.42	51.95	145.95
		2	729.95	3.95	42.61
		3	272.03	1220.29	1578.69
		4	952.74	13.57	401.57
	9	1	1489.34	1197.76	951.10
		2	2969.62	95.46	3762.12
		3	2224.53	256.19	2989.52
		4	2133.63	361.39	1971.72
Wilcoxon Signed Ranked Sum Tests			↑ _____ ↑	↑ _____ ↑	
			↓	↓	
B-series Sessions			36	35	
			0.012(RE)	0.017(MM)	
C-series Sessions			73	75	
			0.008(RE)	0.005(RE)	
All Sessions			205	139	
			<0.001(RE)	0.204(RE)	

Reported metrics for each session are based on the last period in which the state was X, the last period in which the state was Y and the last period in which the state was Z, as applicable. Each statistical test compares RE to an alternative model. Each cell in the lower portion of the table contains the associated test statistic and level of significance as well. The model in parentheses is the one favored by the data based on the median. This table is comparable to Table V in P&S.

Figure C1a. Prices in Market 4 Sessions – Asset X



Period 1 2 3 4 5 6 7 8 9

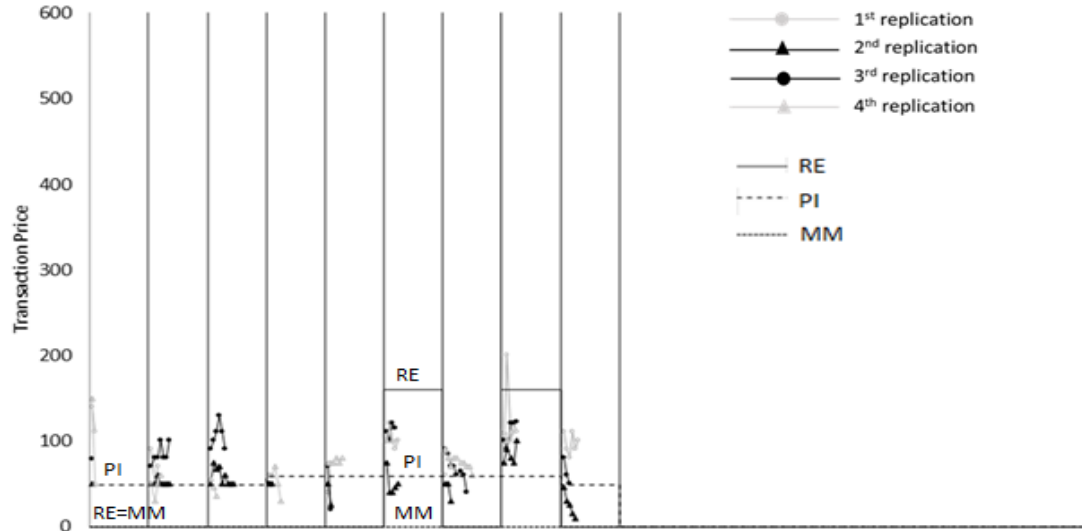
State Z Z Z X X Y X Y Z

Average Prices

P&S 97 10 20 170 215 80 215 - -

Rep 84 52 57 88 107 73 98 69 78

Figure C1b. Prices in Market 4 Sessions – Asset Y



Period	1	2	3	4	5	6	7	8	9
State	Z	Z	Z	X	X	Y	X	Y	Z

Average Prices

P&S	58	57	-	2	-	154	1	158	-
Rep	97	64	72	53	58	89	67	107	64

Figure C1c. Prices in Market 4 Sessions – Asset Z

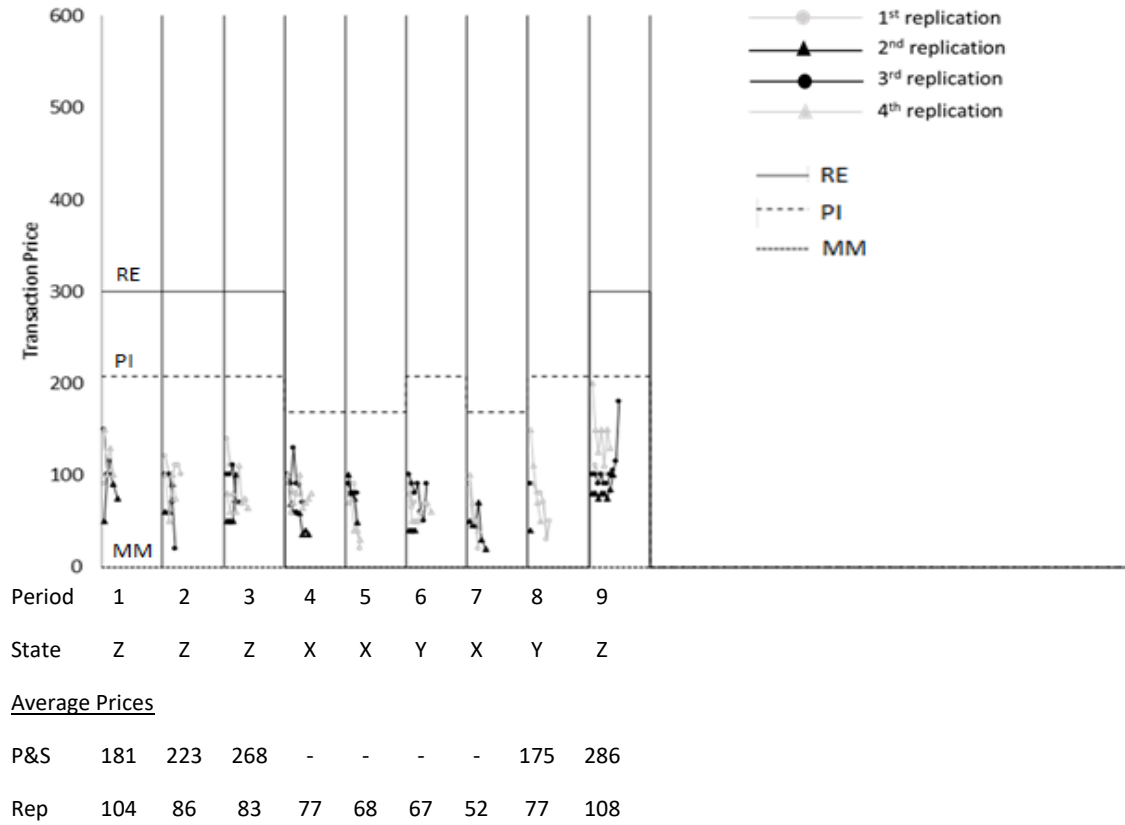
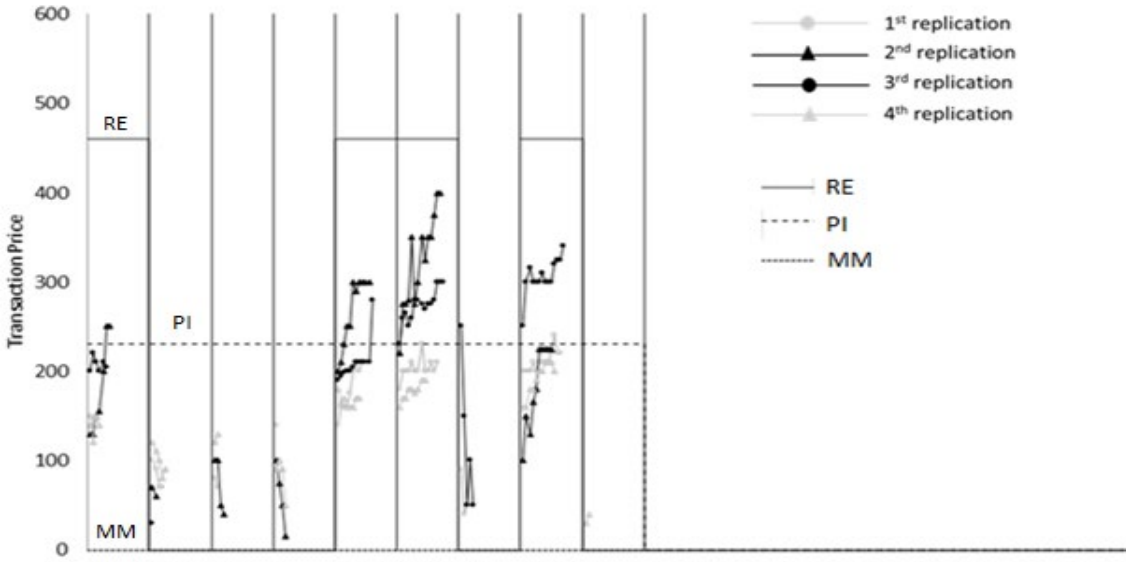


Figure C2a. Prices in Market 5 Sessions - Asset X

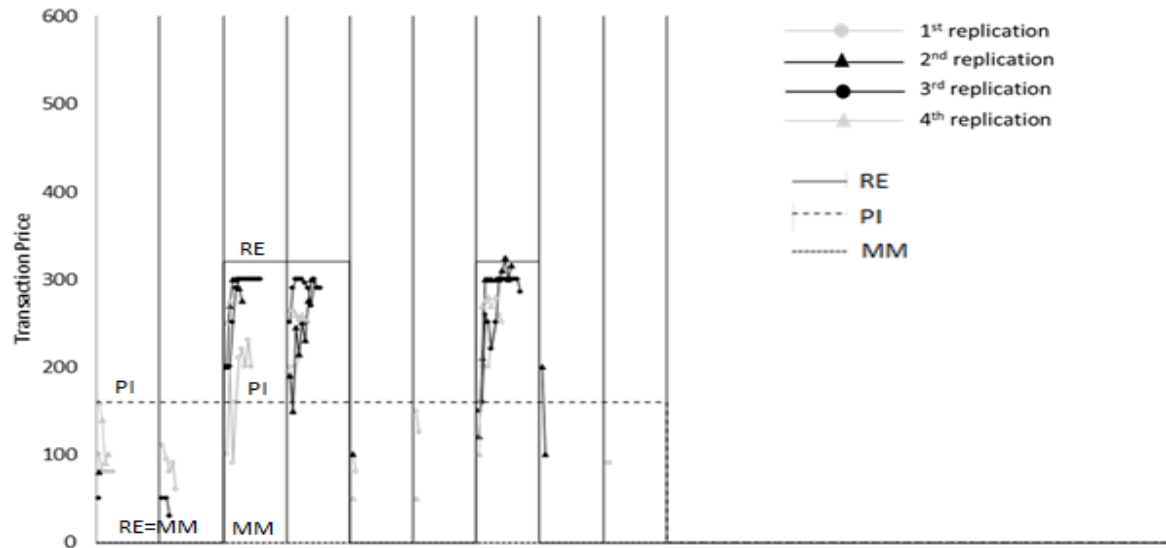


Period	1	2	3	4	5	6	7	8	9
State	X	Z	Y	Y	X	X	Y	X	Z

Average Prices

P&S	181	100	46	4	250	385	-	414	-
Rep	175	83	86	79	210	252	98	229	35

Figure C2b. Prices in Market 5 Sessions - Asset Y

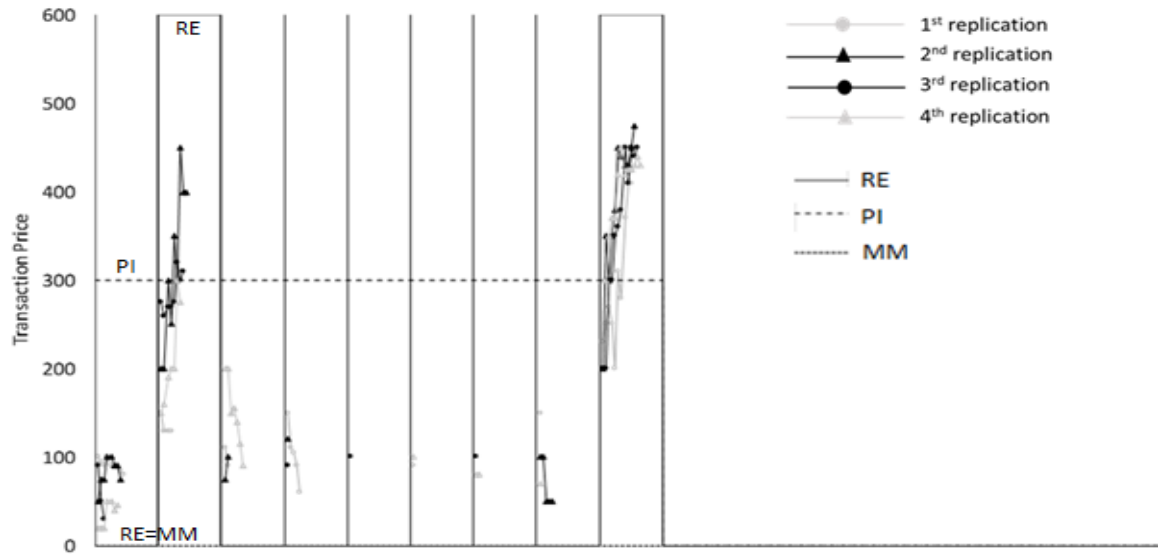


Period	1	2	3	4	5	6	7	8	9
State	X	Z	Y	Y	X	X	Y	X	Z

Average Prices

P&S	156	109	299	296	1	31	306	-	-
Rep	95	71	245	253	83	108	256	150	90

Figure C2c. Prices in Market 5 Sessions - Asset Z

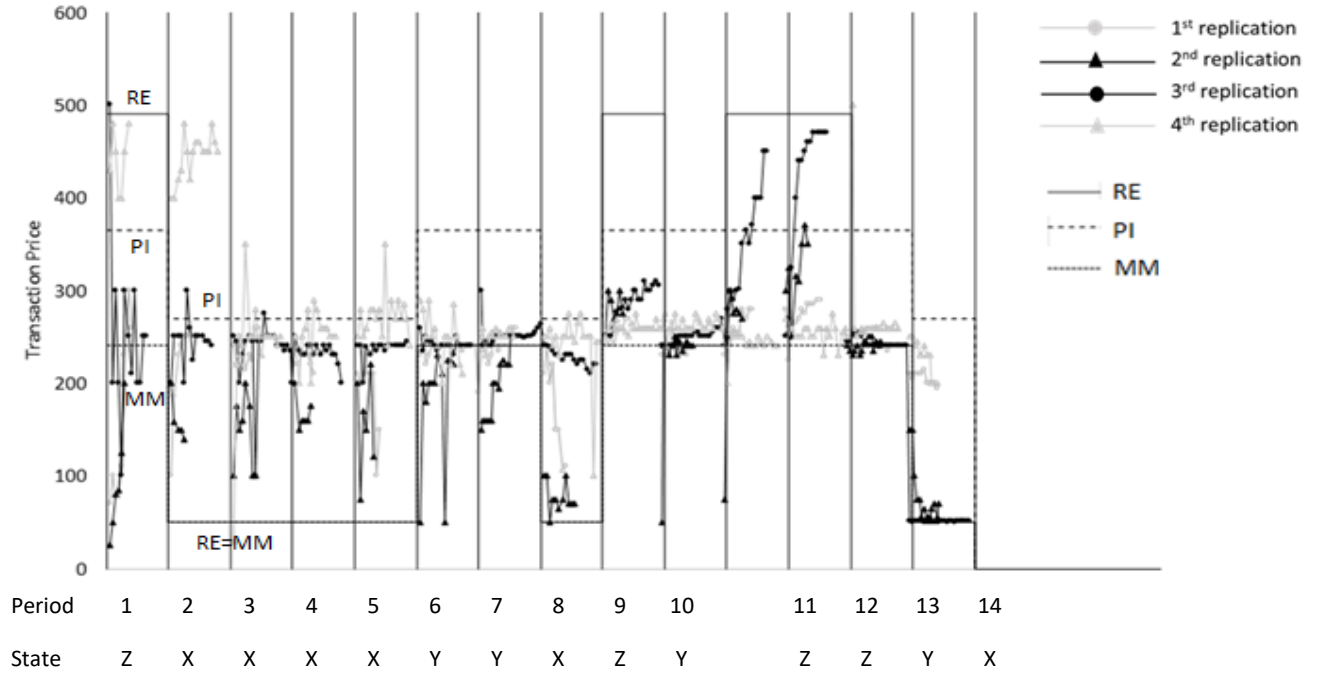


Period	1	2	3	4	5	6	7	8	9
State	X	Z	Y	Y	X	X	Y	X	Z

Average Prices

P&S	193	315	1	67	1	32	-	1	449
Rep	70	255	130	104	100	95	87	81	362

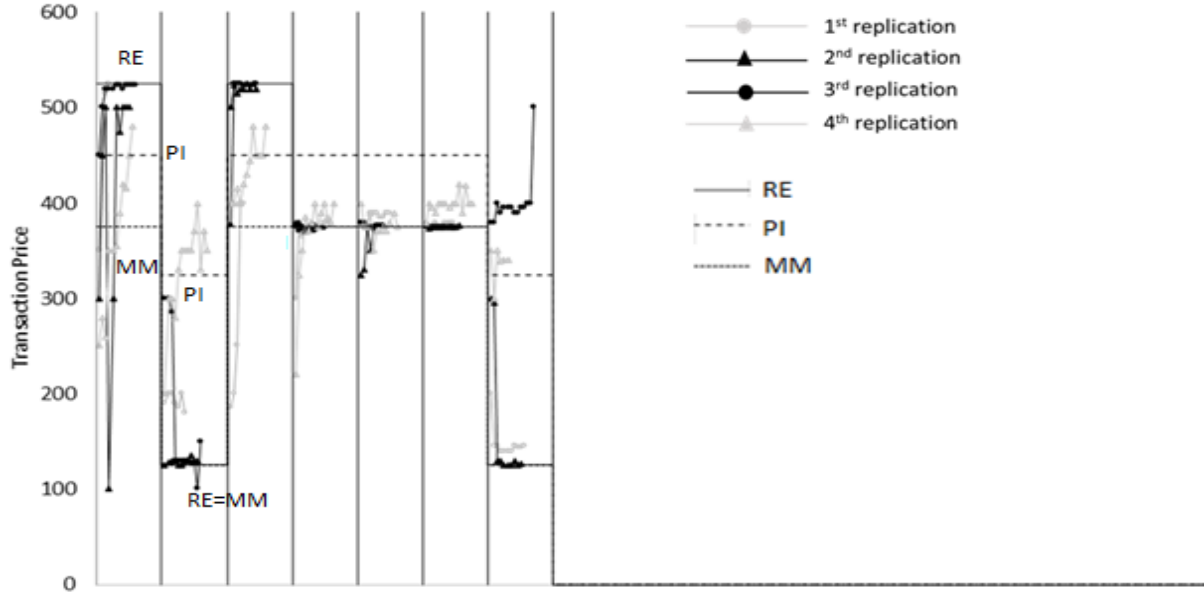
Figure C3. Prices in Market 7 Sessions



Average Prices

P&S	291	222	197	228	170	235	241	148	256	246	348	449	240	84
Rep	239	304	223	229	233	231	238	196	271	248	279	311	249	121

Figure C4. Prices in Market 8 Sessions



Period	1	2	3	4	5	6	7
State	Z	X	Z	Y	Y	Y	X

Average Prices

P&S	512	140	517	376	366	369	126
Rep	438	214	454	370	376	384	265

Appendix D. Additional Findings for Experienced Traders in Series B and C Markets

Tables D1, D2 and D3 compare RE with PI and MM using only data from experienced traders (Markets 5 and 8) for each of the criterion used in P&S.

Table D1. Comparison of Actual Prices to Prices Predicted by Three Models at the End of Each Market for Sessions with Experienced Traders (Markets 5 and 8)

	<u>MAD</u>			<u>Log Odds Under Normality</u>			<u>Percentage of Convergent Price Changes</u>		
	PI	RE	MM	PI	RE	MM	PI	RE	MM
Wilcoxon Signed Ranked Sum Tests	↑ _____ ↑	↑ _____ ↑		↑ _____ ↑	↑ _____ ↑		↑ _____ ↑	↑ _____ ↑	
	↓	↓		↓	↓		↓	↓	
All Sessions	29	25		23	30		30	26	
	0.124(PI)	0.327(RE)		0.484(PI)	0.093(RE)		0.079(RE)	0.262(RE)	

Reported metrics for each session are based on the last period in which the state was X, the last period in which the state was Y and the last period in which the state was Z. Each statistical test compares RE to an alternative model. Each cell in the lower portion of the table contains the associated test statistic and level of significance as well. The model in parentheses is the one favored by the data based on the median. This table is comparable to Table III in P&S.

Table D2. Comparison of Actual Allocations with the Allocations Predicted by Three Models
 Criterion: Percent of Predicted Flow of Securities that Actually Occurred at the End of Each
 Market for Sessions with Experienced Traders (Markets 5 and 8)

	Models		
	PI	RE	MM
Wilcoxon Rank Sum Tests	↑ _____ ↑	↑ _____ ↑	↑ _____ ↑
	↓		↓
All Sessions	39		24
	0.027(RE)		0.083(MM)

Reported metrics for each session are based on the last period in which the state was X, the last period in which the state was Y and the last period in which the state was Z, as applicable. Each statistical test compares RE to an alternative model. Each cell in the lower portion of the table contains the associated test statistic and level of significance as well. The model in parentheses is the one favored by the data based on the median. This table is comparable to Table IV in P&S. However, P&S report that a Wilcoxon Signed Rank Sum Test is used, but that test is not defined for any comparison between RE and another model except for a test between PI and RE in Series-B Markets. The test statistic for such a test would be 0 based on our data with a p-value of 0.012 with the data favoring RE.

Table D3. Comparison of Actual Distribution of Profits to Distribution Predicted by Three
 Models
 Criterion: Squared Sum of Deviations from the Mean Across Investors at the End of Each Market
 for Sessions with Experienced Traders (Markets 5 and 8)

	Models		
	PI	RE	MM
Wilcoxon Signed Ranked Sum Tests	↑ _____ ↑	↑ _____ ↑	↑ _____ ↑
	↓		↓
All Sessions	32		19
	0.050(RE)		0.889(RE)

Reported metrics for each session are based on the last period in which the state was X, the last period in which the state was Y and the last period in which the state was Z, as applicable. Each statistical test compares RE to an alternative model. Each cell in the lower portion of the table contains the associated test statistic and level of significance as well. The model in parentheses is the one favored by the data based on the median. This table is comparable to Table V in P&S.

Appendix E: Additional Evidence on Series D for Markets 12 and 13

This appendix compares RE to PI and MM on the basis of profit, as was done in Table C2 for Series B and C markets. The analysis is presented in Table E1. For Series D, neither RE nor MM make a prediction regarding who should hold the assets and thus, no analysis analogous to that reported in Table C1 is possible.

Table E1. Comparison of Actual Distribution of Profits to Distribution Predicted by Three Models
Criterion: Squared Sum of Deviations from the Mean Across Investors at the End of Each Market

<u>Market Experiments</u>			<u>Models</u>		
Series	Market	Replicate	PI	RE	MM
D	12	1	632.54	45.88	45.88
		2	980.33	0.00	0.00
		3	1205.90	70.84	70.84
		4	846.41	13.81	13.81
	13	1	2801.49	16.40	16.40
		2	3265.36	1.41	1.41
		3	3213.30	0.57	0.57
		4	3242.76	0.00	0.00
Wilcoxon Signed Ranked Sum Tests			↑	↑	↑
			↓		↓
D-series Sessions			36		
			0.012(RE)		
B&D Sessions			136		
			<0.001(RE)		
C&D Sessions			183	152	
			<0.001(RE)	0.003(RE)	
All Sessions			399	235	
			<0.001(RE)	0.248(RE)	

Reported metrics for each session are based on the last period in which the state was X, the last period in which the state was Y and the last period in which the state was Z, as applicable. Each statistical test compares RE to an alternative model. Each cell in the lower portion of the table contains the associated test statistic and level of significance as well. The model in parentheses is the one favored by the data based on the median. This table is comparable to Table V in P&S.

Appendix F: Alternative Estimates of PSP

Since effect sizes may be overestimates in small-sample studies, we repeat the analysis reported in Tables 4 and 8 assuming that the actual effect size is two thirds of the implied effect size given the p-values reported in P&S. Under this assumption, the power attained by P&S to detect such an effect size in Series B and C between RE and PI was 0.71 and between RE and MM was 0.6263. The implied power of our replication under this assumption is 0.99983 and 0.9987, respectively. As seen in Table F1, the implications are qualitatively unchanged. In light of the additional evidence provided by our replication, the PSP that RE outperforms the other models in Series B and C markets is quite low, unless one had an extremely strong prior.

Table F1. Post Study Probability that MAD Lower Under RE than PI or MM in Series B and C

<u>Probability RE outperforms PI</u>			<u>Probability RE outperforms MM</u>		
Prior	PSP after P&S	PSP after Replication	Prior	PSP after P&S	PSP after Replication
0.99	0.9993	0.2056	0.99	0.9992	0.6371
0.95	0.9963	0.0437	0.95	0.9958	0.2520
0.9	0.9923	0.0230	0.9	0.9912	0.1376
0.75	0.9771	0.0078	0.75	0.9741	0.0505
0.5	0.9343	0.0026	0.5	0.9261	0.0174
0.25	0.8259	0.0009	0.25	0.8068	0.0059
0.1	0.6126	0.0003	0.1	0.5819	0.0020
0.05	0.4282	0.0001	0.05	0.3973	0.0009
0.01	0.1257	<0.0001	0.01	0.1123	0.0002

Under the same assumption, our power to detect two thirds of the reported effect size underlying Table 8 (i.e. an effect of $z = 1.75$) is 0.9791. Table F2 shows that under these assumptions the results in Table 8 are also robust.

Table F2. Post Study Probability that MAD Lower Under RE than PI or MM for Series D

<u>Probability RE outperforms PI</u>		<u>Probability RE outperforms MM</u>	
Prior	PSP	Prior	PSP
0.99	0.9995	0.99	0.9995
0.95	0.9973	0.95	0.9973
0.9	0.9944	0.9	0.9944
0.75	0.9833	0.75	0.9833
0.5	0.9514	0.5	0.9514
0.25	0.8672	0.25	0.8672
0.1	0.6851	0.1	0.6851
0.05	0.5076	0.05	0.5076
0.01	0.1651	0.01	0.1651