Uncertainty, Anchoring, and Expectations Formation: Experimental Evidence[1]

Benjamin Radoc
Bangko Sentral ng Pilipinas

Sarah Lynne S. Daway-Ducanes
University of the Philippines

[Preliminary draft]
August 2024

Abstract

The important role of expectations in intertemporal decision making has long been recognized but disagreement among economists on how expectations are formed persists. We conducted an online learning to forecast experiment to determine the impact of uncertainty (in terms of group size and market volatility) and anchoring (or a non-binding target price band) on the quality of price forecasts. Controlling for learning and order effects, we find that forecast errors are significantly higher in more volatile markets but not in larger groups. The reduced forecast errors in later markets confirm the importance of learning and experience, but there is mixed evidence supporting the rational expectations hypothesis even with the aid of an anchor.

Keywords: expectations formation, laboratory experiment, bounded rationality

JEL Classification: C91, E31, E71

[1] We are grateful to the seminar participants of the BSP Research Huddle for their comments and suggestions. We also thank the Philippine Competition Commission for generously funding this research, and to Juan Felipe Coronel for programming assistance. The protocol on data collection was cleared by the Ateneo de Manila University Research Ethics Office (Protocol ID No. AdMUREC_22_043). All errors are the responsibility of the authors. Corresponding author: radocbe@bsp.gov.ph.

1. Introduction

Since economic agents commonly make intertemporal decisions, the important role of expectations in decision making has long been recognized. Expectations influence individual choice which in turn determine macroeconomic outcomes (Colasante, et al., 2017). However, disagreement among economists on *how* (unobservable) expectations are formed persists as the nature of the information and how these are used by agents are unclear (Cavallo, et al., 2017). As such, macroeconomists typically specify *a priori* how expectations are formed (Brayton, et al., 1997). Under the *rational expectations hypothesis* or REH (Muth, 1961), it is assumed that agents quickly learn from past mistakes, use all available information when deciding, and understand the underlying mechanism that leads to an outcome. The full information rational expectations assumption, as a workhorse to model the expectation formation process, suggests that forecast errors are unpredictable. Meanwhile, under the *adaptive expectations* or *error learning hypothesis* (Burmeister & Turnovsky, 1976), agents adjust forecasts conditional on past-period forecast error. They slowly change their expectations even with new information, and past trends are used to predict future outcomes. Since there is a gradual change in expectations, forecast errors are correlated with past errors. In recent years, *diagnostic expectations* (Gennaioli & Shleifer, 2018) have become a salient departure from the REH. When agents are susceptible to the *representativeness heuristic* (Kahneman & Tversky, 1972), they overestimate the probability that a past event will recur, extrapolate that event into the future, and tend to overreact (L'Huillier, et al., 2023). The resulting price path is characterized by an initial underreaction, followed by overshooting, and then a crash (Bordalo, et al., 2021).[2]

Incorporating *bounded rationality* (or the recognition that agents typically lack access to full information or have limited time and cognitive resources) into the expectation formation process has generated alternative models with incomplete information or agents with imperfect response to new information. *Information frictions* lead to inattention (e.g., ignoring inflation especially in societies where a low inflation environment persists) or reliance on personal experience despite being an inaccurate information source (Cavallo, et al., 2017). Information frictions may also arise when agents are uncertain whether a signal represents either relevant new information or noise (Coibion, et al., 2018). *Sticky information* in terms of the speed of information transmission or the ability of agents to process new information (Mankiw & Reis, 2002) can result in heterogeneous expectations. Differences in demographic characteristics and experience can also explain the heterogeneity in expectations (Malmendier & Nagel, 2015). Meanwhile, *information acquisition* or the manner in how the new information is gathered (Armantier, et al., 2016) also matters. For example, learning by description

[2] Frydman & Frydman (2022) argued that the tendency of decision makers to overreact to information is an artifact of the diagnostic expectations hypothesis.

(when the information is presented) and learning by experience (when the information is gathered from experience) can result in differences in the updating of beliefs (Hertwig et al., 2004).[3]

The complexity of having multiple theories on expectations formation is aggravated by data issues. Expectations surveys[4] allow representative sampling, but the data collected may be inaccurate if the instrument is not incentive compatible (i.e., when a fixed amount is paid to all respondents regardless of the quality of responses). Alternatively, expectations may be inferred from financial instruments, but pricing may be confounded by other factors such as liquidity and risk premia (Cornand & Hubert, 2020).

An economic experiment is an alternative method to collect data on decision markers' expectations.[5] In *learning to forecast experiments* (LtFE), researchers can elicit expectations directly in a controlled environment (Marimon & Sunder 1994). Over several periods, participants learn to predict the price of an asset without explicit knowledge of the underlying market equilibrium equation. In asset pricing tasks, participants know the dividends and observe the past realized prices so it is possible to compute the fundamental price of an asset (Hommes et al., 2008). Since participants are compensated depending on the quality of their predictions, they have a strong incentive to form rational expectations.

Bao, et al. (2021) summarized trends in LtFEs published after 2010. (i) Convergence to REH is faster under LtFEs than in *learning to optimize experiments*.[6] (ii) Bubbles and crashes are persistent in small markets (6 to 10 participants) and in large-scale LtFEs. (iii) Markets are more stable when participants make long-run expectations (versus one-period or two-period ahead forecasts). (iv) Predicting returns is more susceptible to bubbles than predicting prices. (v) In performing complex tasks, participants perform better when heuristics are used. (vi) The forecasts of participants with higher cognitive abilities converge to REH, but experience also matters. (vii) There is consistency in the results from the lab and in the field. (viii) LtFEs are increasingly useful in examining the role of monetary policy (i.e., interest rates, central bank communication) in stabilizing prices.

---

[3] In a recent paper, Wang (2024) compared the dispersion of density forecasts (or uncertainty) under sticky expectations, noisy information, and diagnostic expectations. Uncertainty is higher with sticky expectations due to the lagged updating of information. With noise, the additional uncertainty arises from variations agents' reaction to new information. Under diagnostic expectations, uncertainty is comparable to full information rational expectations and assumes a mean-reversion of overreaction. Overall, uncertainty is larger if diagnostic expectations are combined with noisy information.

[4] A fair test of the REH requires conditions that survey datasets cannot satisfy. As such, directly rejecting REH based on survey data is difficult. However, openness on the use of survey data has changed amid recent crises and new theories on expectations formation (Born, et al., 2023).

[5] The controlled environment in experiments has been criticized as potentially limiting the external validity of experiments. Cornand & Hubert (2022) compared experimental data with inflation forecasts by different agents (professional forecasters, industrial forecasters, central bankers, households, financial market participants). Although there is considerable heterogeneity across the data sources, common features emerged: large forecast errors, autocorrelated forecast errors, and predictable forecast revisions. Experimental data also exhibit similar biases.

[6] In a *learning to optimize experiment*, group participants decide on a production quantity. The resulting market price is based on the total supply and earnings correspond to individual profits (Bao, et al., 2021).

Against this backdrop, we conducted an LtFE to examine the pattern of forecast errors conditional on the level of induced uncertainty and the presence of a non-binding target price band. In our between-subjects design with groups, uncertainty is induced by rolling either a 3-sided or an 8-sided die, and by varying group size (3 players or 6 players). Participants know that actual price x_{t+1} is determined by a stochastic difference equation but they do not know the equation. x_{t+1} for each group is a function of x_t, the group's median forecast error at period t, and a random number u_t. By design, predicting x_{t+1} is more difficult when u_{t+1} is high (i.e., when an 8-sided die is rolled) and when group size is large.[7] Meanwhile, the non-binding target price in the experiment period is a simplified inflation target for each period. Instead of the usual inflation target range (in percent), we show a series of paired low and high prices that participants may either ignore or use as an anchor in predicting x_{t+1}.

Controlling for learning and order effects, we find that forecast errors are consistently higher in more volatile markets but evidence on group size effect is mixed. The significant reductions in forecast errors in later markets confirm the importance of learning or forecasting practice, but there is mixed evidence that participants form rational expectations even with a non-binding target price range as a possible anchor. This result offers important implications on communicating monetary policy.

The rest of the paper is organized as follows. Section 2 describes the experimental design, Section 3 presents the results, and Section 4 concludes.

2. Experimental Design

Our LtFE is designed to determine the impact of group size g, volatility u, and a non-binding target price band on the pattern of price forecasts. We hypothesize that with experience, forecast errors are more likely to converge towards REH under low uncertainty (when $g = 3$ and/or $u_t \in (1, \ldots, 3)$) than under high uncertainty (when $g = 6$ and/or $u_t \in (1, \ldots, 8)$). Meanwhile, the introduction of an anchor (in the form of a target price range) facilitates convergence towards REH.

The online experiment is programmed in *Python* and the user interface is implemented in *o-tree* (Chen, et al., 2016) to allow participants to access the experiment through a mobile gadget. Each participant i monitors price x_t of a homogeneous commodity over two markets with thirty periods each. The known range for x_t is 1 to 500 denominated as experimental currency (ECU). A participant's task is to predict x_{t+1} by stating her own forecast f_{t+1}. Payoff in each period is inversely related to the forecast error, $x_{t+1} - f_{t+1}^{i}$.

Following Becker, et al. (2007), x_{t+1} is derived from this simplified stochastic difference equation with a positive feedback mechanism, i.e., higher forecasts lead to higher actual prices (Bao, et al., 2021):

[7] Across treatments, the range of median forecast errors in groups of 6 (-301 to +31) is wider than in groups of 3 (-440 to +117).

$$x_{t+1} = x_t + int\left(\frac{1}{10} \cdot mdn\left(x_t - f_t^i\right)\right) + u_{t+1}. \quad (1)$$

x_{t+1} is a function of the current price x_t, the group's median forecast error $(x_t - f_t^i)$, and a random number u_{t+1}. The function *int* ensures that all values of the time series are integers. This formulation allows aggregate expectations to affect the actual price similar to how expectations influence macroeconomic dynamics (Rholes & Petersen, 2021).

At the start of an experiment session[8], either groups of 3 players or 6 players are randomly formed among the online participants. A session has two markets with 30 periods each: a low-volatility market (L) and high-volatility (H) environment. In the low-volatility environment, u_t is randomly determined from the roll of a 3-sided die, $u_t \in (1, 2, 3)$. To induce a high-volatility setting, an 8-sided die is used, $u_t \in (1, \dots, 8)$. The order of L and H markets is reverse in some treatments to capture possible order effects.

In period 1 of market 1, participants do not know that the value of x_t is randomly drawn from a range of ECU 60 to 70. A similar procedure is done for market 2. Although the tasks in markets 1 and 2 are similar, participants know that the two markets are independent, i.e., although the unknown stochastic difference equation is the same, the outcomes in market 1 do not carry over to market 2.

For the treatments with a target price band, participants are shown two diagonal lines with a height of ECU 30 (starting at $x_1 \pm 15$) to indicate the target price range at each period. Since the price range is non-binding (i.e., there is no penalty for breaching the upper or lower bound), participants can choose to either ignore or use it as an *anchor* when predicting x_{t+1}. An anchor is a non-informative numerical cue (e.g., price): in this experiment, the target price range is not considered in the determination of x_{t+1} so that a Bayesian decision maker may opt to ignore this cue.

The six experiment conditions are summarized in Table 1. Since not all planned sessions with 18 to 21 target participants were filled, we conducted a total of 72 sessions from September to October 2023.[9] Each session lasted for an average of 80 minutes. All participants are students from various colleges and universities in the Philippines. Age range is 18 to 46 years. 62% of the participants are female.

Each period, participants are reminded onscreen that they have 30 seconds to encode their price prediction. A period ended only when all members of a group have inputted a forecast. To indicate forecasting accuracy per period, players are shown a historical line graph of their own past forecast, the actual price, and the non-binding target price band (if applicable).

---

[8] On the day of the experiment, each registered participant received via email a unique link for the experiment and a *Zoom* meeting link where participants had masked identities. At the designated time, all session participants logged into the *Zoom* meeting. Laboratory assistants monitored the *Zoom* meeting so they can privately respond to participants who had a question on the experiment instructions or had technical problems.

[9] Since the authors did not have access to a web-based recruitment system such as ORSEE (Online Recruitment System for Economic Experiments), a recruitment poster was shared online to invite university students in the Philippines to register.

Table 1. Summary of experiment conditions

Group size	Market 1 volatility	Market 2 volatility	Target range displayed?	Number of groups	Number of sessions	N
3	Low	High	No	40	16	120
3	High	Low	No	41	13	123
6	Low	High	No	20	10	120
6	High	Low	No	20	10	120
6	Low	High	Yes	21	12	126
6	High	Low	Yes	20	11	120
Total					72	729

In each market, participants start with an individual endowment of ECU 1,500. After each period, the ECU equivalent of a participant's prediction error is deducted from the remaining endowment (Table 2). We put a cap on the deductions to ensure that no participant ends with negative earnings after the last period.

Table 2. Range of prediction errors and corresponding deductions

Deductions (in ECU)	Prediction error
0	0
-5	1-3
-10	4-6
-15	7-9
-20	10-12
-25	≥ 13

The total remaining ECU in markets 1 and 2 are multiplied by 0.25 to determine the Philippine peso equivalent of the earnings. The maximum earnings are PHP 750 plus a show-up fee of PHP 100. Average actual earnings were PHP 358.82 (USD 6.52) with a standard deviation of PHP 29.66 (USD 0.54). Earnings were transferred to a nominated digital wallet (*GCash* or *Maya*) within 36 hours after the experiment. Refer to *Annex 1* for the sample experiment instructions, *Annex 2* for a sample experiment computer screen, and *Annex 3* for the post-experiment questionnaire.

3. Results

We analyzed participants' prediction errors to determine the effect of volatility u_t and group size g, and a non-binding target price range on the quality of price forecasts. Where appropriate, we examined *raw* forecast errors calculated as $(forecast - actual\ price)$ and *absolute value* of forecast errors computed using this formula: $\sqrt{(forecast - actual)^2}$ disaggregated between markets 1 and 2.

Figure 1 depicts the distribution of per round forecast errors across the experiment conditions where each bar represents a bucket of 5 periods. Although participants know that market 1 and market 2 are independent, they are also aware that the unknown stochastic difference equation used is the same, except for range of values for u_t or the die rolled to determine the random number. The results comparing forecast errors in markets 1 and 2 point to the importance of learning or forecasting practice. Regardless of the order of L and H markets, forecast errors in market 2 are significantly lower than in market 1, whether in 3-player sessions (z=-3.001, p=0.0027), 6-player sessions (z=-3.878, p=0.0001), or 6-player sessions with a target price range (z=-3.724, p=0.0002). Comparisons of raw forecast errors and the absolute value of forecast errors conditional on the order by which participants encountered L and H markets do not indicate any order effect (see *Annex 4*).

[INSERT FIGURE 1]

Result 1: Controlling for learning and order effects, forecast errors are consistently higher in more volatile markets but evidence on group size effect is mixed.

Figure 2 shows the distribution of actual prices and forecasts across the 6 experiment conditions, disaggregated between L and H markets. To determine a group size effect while controlling for volatility u_t and any order or learning effect, we compare the forecast errors of market 1 in Figures 2(a) and 2(c), and market 1 in Figures 2(b) and 2(d). The Wilcoxon (Mann-Whitney) ranksum test results indicate a statistically significant difference in the forecast errors between 3-player groups and 6-player groups in high-volatility markets (z=-2.234, p=0.0255) but not in low-volatility markets (z=-0.198, p=0.8433). The pattern in the results is consistent when we analyze absolute forecast errors in high-volatility markets (z=3.862, p=0.0001) and low-volatility markets (z=0.159, p=0.8433).

[INSERT FIGURE 1]

We also examined the impact of u_t on forecast errors while controlling for group size and order or learning effects. Actual prices in market 1 of Figure 2(a) (SD=17.83) are significantly less dispersed than in Figure 2(b) (SD=40.31). A similar pattern holds for market 1 in Figure 2(c) (SD=18.09) and Figure 2(d) (SD=39.58). Among 3-player groups, forecast errors are significantly larger in high-volatility markets relative to low-volatility markets (z=11.321, p=0.0000). A similar pattern holds among 6-player groups (z=4.687, p=0.0000).

Result 2: The general pattern in the forecast errors suggests learning. Indications that participants form rational expectations is mixed.

Given a positive feedback mechanism in equation (1), there is an upward trend in actual prices and forecasts across all experiment conditions, with an expected steeper slope for H markets. For example, market 1 in panel (a) shows that average actual prices in rounds 26-30 are 73.5% higher than in rounds 1-5, while average forecast prices are higher by 24.2%. Market 2 in panel (a) indicates an average increase in average actual prices and forecast prices at 158.9% and 140.4%, respectively.

The table under each panel in Figure 2 summarizes the raw and absolute value of forecast errors in markets 1 and 2. The significantly lower average raw prediction errors and absolute value prediction errors in market 2 of panels (a) to (d), along with lower standard deviations across experiment conditions, indicate that participants learned to improve the quality of their forecasts.[10] In panel (e) where a non-binding target price band is available, the reduction in raw forecast errors between markets 1 and 2 is less pronounced but statistically significant (z=8.17, p=0.000). A similar result holds when we compare the absolute value of forecast errors (z=-25.45, p=0.000).

To visualize the distribution of forecast errors and validate the results observed in Figure 1, we plotted the quantiles on actual prices and forecasts in Figure 3. The Q-Q plots corroborate the observed improvement in forecast accuracy.

[INSERT FIGURE 3]

Given the better quality of predictions in market 2, does the pattern in the forecast errors approximate rational expectations? REH predicts that prediction errors are unpredictable or uncorrelated. To confirm this, we examined the autocorrelation of forecast errors, by treatment and market volatility.

An autocorrelation test describes the degree of correlation between a variable and its past values. Table 3 summarizes the Box-Pierce Q statistics up to *t-3*. (refer to *Annex 4* for more details). As shown, we reject the null hypothesis that the correlation up to lag *k* is equal to zero, except for the condition where participants encountered a target price range (last row). This suggests that participants' process of forming expectations is generally not consistent with REH. However, there are instances where participants form expectations close to REH: when they learn to forecast in a more volatile market but with the aid of a target price band (market 1), followed by a less volatile environment (market 2).

[10] Participants' ability to learn to forecast as they gain experience has been demonstrated in past studies. For example, Armantier, et al. (2016) provided two types of inflation-relevant information to respondents: (i) past-year average food price inflation, or (ii) the average forecast of next-year inflation. Respondents updated their inflation expectations sensibly (e.g., revised down if they overestimated) and were more receptive to new information under greater inflation uncertainty.

Table 3. Box-Pierce Q statistics

	lag(1)	lag(2)	lag(3)
3-players, L-H: market 1 (L)	255.83***	466.67***	560.74***
3-players, L-H: market 2 (H)	1529.70***	2702.00***	3693.00***
3-players, H-L: market 1 (H)	382.48***	573.60***	654.23***
3-players, H-L: market 2 (L)	7.06***	7.35 **	7.37 *
6-players, L-H: market 1 (L)	349.27***	470.98***	526.12***
6-players, L-H: market 2 (H)	4.48 **	4.65 *	**4.69**
6-players, H-L: market 1 (H)	835.31***	1332.2***	1828.7***
6-players, H-L: market 2 (L)	1871.10***	3491.40***	4974.00***
price band, L-H: market 1 (L)	22.42***	24.62***	44.52***
price band, L-H: market 2 (H)	64.54***	171.8***	203.28***
price band, H-L: market 1 (H)	303.05***	369.43***	371.82***
price band, H-L: market 2 (L)	**1.61**	**1.72**	**2.64**

Note: A simplified version of the Ljung-Box test, this measure is used to examine residuals from a time series to determine if all error autocorrelations up to a point are equal to zero. The corresponding null hypothesis is that all correlation up to lag *k* are equal to 0. *=p<0.1, **=p<0.05, ***=p<0.01

Result 3: Expectations anchored on a target price range reduces forecast errors, but the anchoring persists only when the anchor remains plausible.

The experiment conditions with a target price range are designed to determine the effect of an *anchor* on forecast errors. Although past experiments showed that valuations are positively correlated with an arbitrary anchor price, the anchoring effect depends on the task, e.g., stronger for selling vs buying, and the type or quality of anchors, e.g., plausible vs implausible (Sugden, et al., 2013).

In this experiment, some participants are shown a continuous diagonal price range (refer to Annex 2) that is independent from equation (1). Since the target price range is *non-binding*, i.e., there is no penalty for breaching the lower or upper bound, participants may either ignore or use it as an anchor when predicting the actual price.

We find that showing a target price range produced significantly lower forecast errors (mean=-0.058) relative to the condition without an explicit anchor (mean=5.048; z=14.296, p=0.0000).[11] This supports past studies showing that clear and simple central bank communication, e.g. a clear inflation target, helps to stabilize expectations (Bao, et al., 2021) and influences less experienced forecasters (Kryvtsov & Petersen, 2021).

[11] This result holds when we compare absolute forecast errors with a target price band (z=-2.101, p=0.0356).

If price expectations are anchored, forecasts should be close to the target price. To measure the degree of anchoring (Czudaj, 2024), we computed the absolute deviation of forecasts from the target as $\sqrt{(forecast - target)^2}$. We considered two anchors: the lower bound and upper bound. Table 4 summarizes the frequency of absolute deviations by experiment condition. A forecast is considered a breach if it is either below the lower bound or above the upper bound. A forecast is anchored if it falls within the lower and upper bounds. [12]

Table 4. Forecast deviations from the target price range (%)

Condition	Within the target price range	Breached upper bound	Breached lower bound
L-H: market 1 (L)	52.54	47.46	0.00
L-H: market 2 (H)	21.11	78.89	0.00
H-L: market 1 (H)	23.75	73.75	2.50
H-L: market 2 (L)	60.67	39.33	0.00

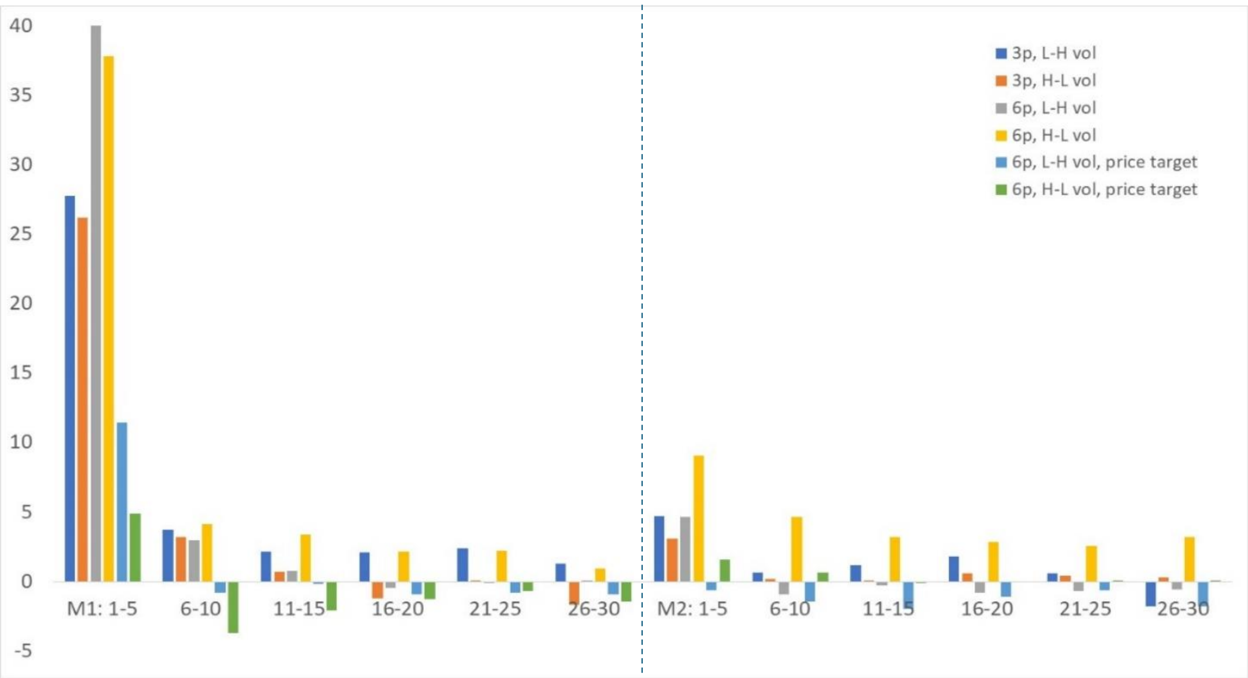
Under low-volatility markets, more than half of forecasts fall within the target price and none of the forecasts are less than the lower bound. Meanwhile, under a high-volatility scenario, less than 25% of forecasts are within the price band and majority of the forecasts breached the upper bound. Given the positive feedback mechanism, ignoring the anchor when the actual price has breached the upper bound (hence, implausible) is not surprising.

4. Conclusion

Learning to forecast experiments are commonly used to study macroeconomic or financial decisions in a controlled environment. We conducted a new LtFE to determine the impact of induced uncertainty and a possible anchor on the quality of price forecasts. Uncertainty is induced by volatility from the roll of a die and by varying group size. We find a robust volatility effect but a weak group size effect. Significantly lower forecast errors in markets with a non-binding target price range suggests anchoring of expectations. However, given the positive feedback mechanism during the experiment, anchoring of expectations persists only when the upper limit of the target price range is plausible. Lower forecast errors in later markets indicate that decision makers can learn to forecast, especially when a target price band is introduced. However, the pattern in forecast errors provides mixed evidence that support the rational expectations hypothesis.

[12] In their LtFE experiment, Rholes & Petersen (2021) showed that 1-period ahead and 2-period ahead forecast errors are significantly lower when participants are given point projections than when density projections are given. However, we are unable to validate that result given our treatment conditions.

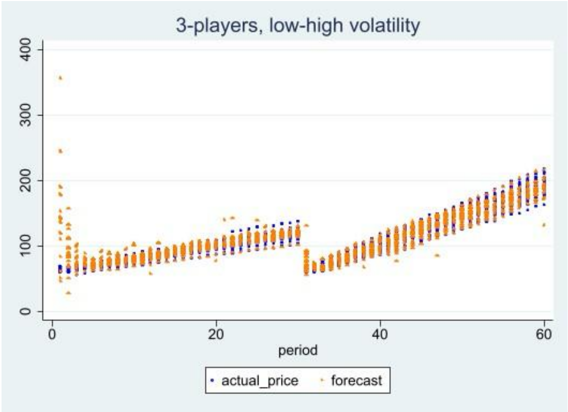
Figure 1. Forecast errors by experiment conditions



Note: Each bar represents a bucket of 5 periods. 3p=group with 3 players; 6p=group with 6 players; L-H vol = low volatility periods followed by high volatility periods; H-L vol = high volatility periods followed by low volatility periods; M1=market 1; M2: market 2.

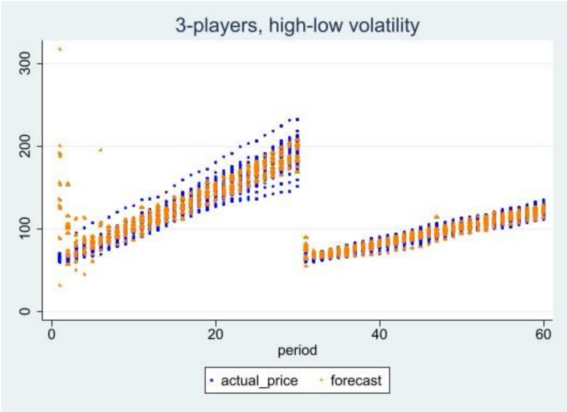
Figure 2. Actual price and forecasts for market 1 (periods 1-30) and market 2 (periods 31-60)

(a)



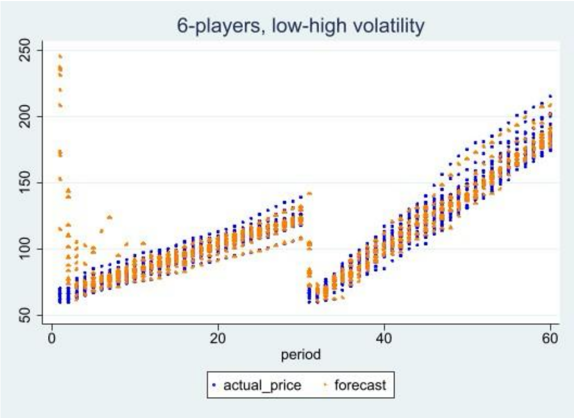
Forecast errors	Market 1 (low volatility)		Market 2 (high volatility)	
	raw value	absolute value	raw value	absolute value
mean	6.58	9.13	1.20	6.05
std dev	41.24	40.75	23.40	22.64

(b)



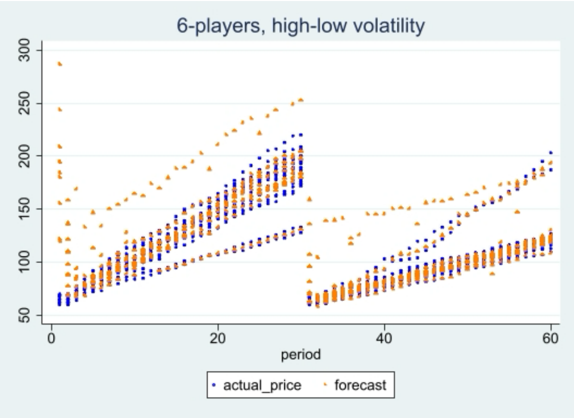
Forecast errors	Market 1 (high volatility)		Market 2 (low volatility)	
	raw value	absolute value	raw value	absolute value
mean	4.55	10.45	0.79	2.23
std dev	40.20	39.09	11.31	11.11

(c)



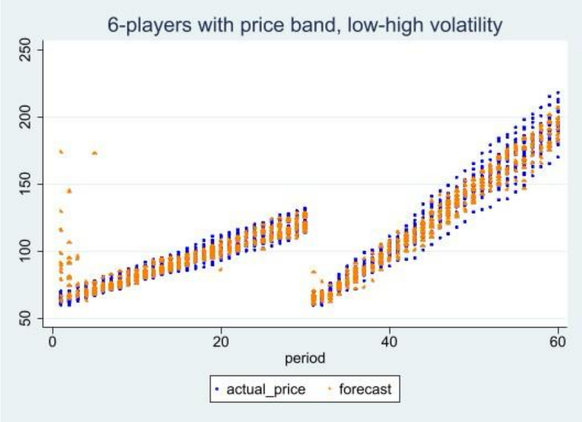
Forecast errors	Market 1 (low volatility)		Market 2 (high volatility)	
	raw value	absolute value	raw value	absolute value
mean	7.22	9.36	0.25	3.70
std dev	43.29	42.88	11.59	10.98

(d)



Forecast errors	Market 1 (high volatility)		Market 2 (low volatility)	
	raw value	absolute value	raw value	absolute value
mean	8.45	13.02	4.27	5.99
std dev	51.39	50.43	38.00	37.77

(e)



Forecast errors	Market 1 (low volatility)		Market 2 (high volatility)	
	raw value	absolute value	raw value	absolute value
mean	1.33	4.27	-1.23	3.41
std dev	23.31	22.95	6.93	6.15

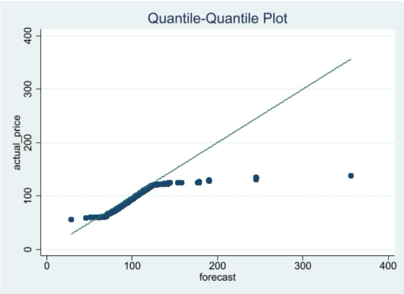
(f)

Forecast errors	Market 1 (high volatility)		Market 2 (low volatility)	
	raw value	absolute value	raw value	absolute value
mean	-0.70	4.98	0.36	1.70
std dev	18.52	17.86	6.37	6.15

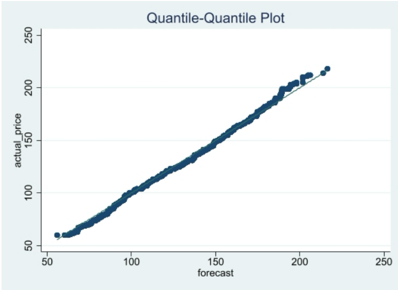
Figure 3. Quantile-Quantile Plot of Actual Prices and Forecasts

3-players, low-high volatility

(a)

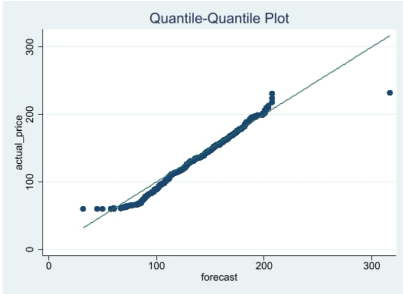


(b)

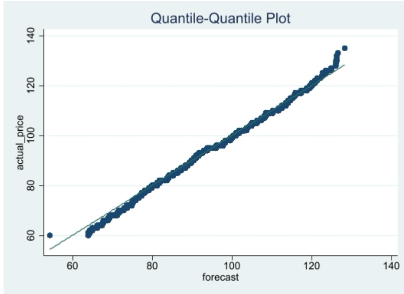


3-players, high-low volatility

(c)

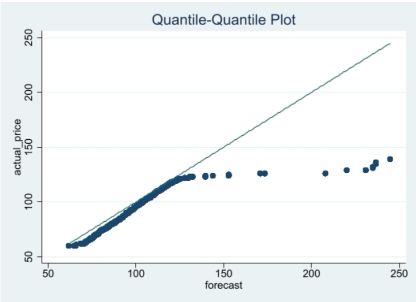


(d)

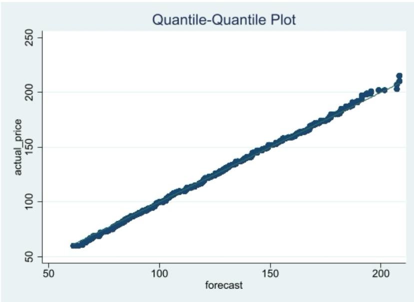


6-players, low-high volatility

(e)

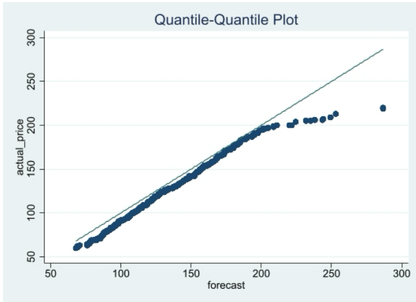


(f)

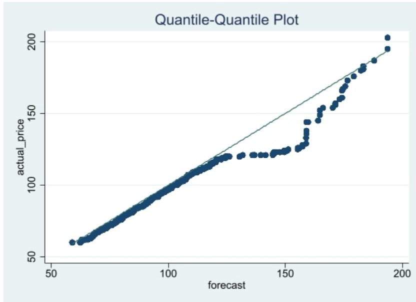


6-players, high-low volatility

(g)



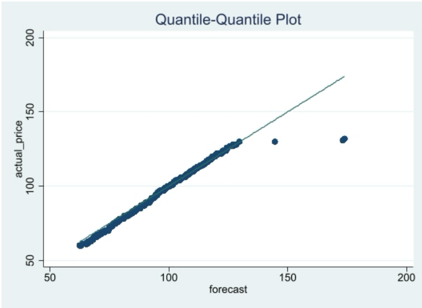
(h)



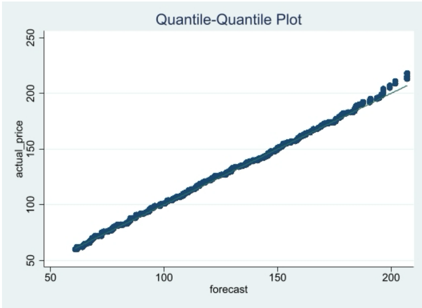
6-players, low-high volatility

6-players, high-low volatility

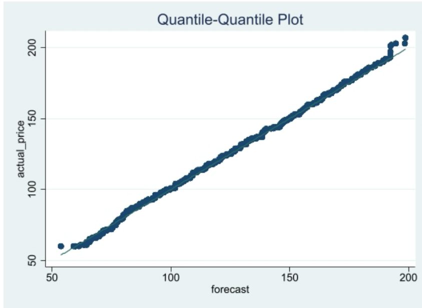
(i)



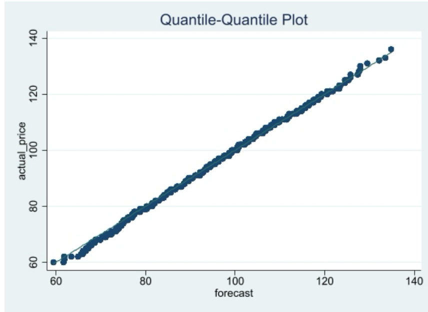
(j)



(k)



(l)



References

Armantier, O., Nelson S., Topa, G., van der Klaauw, W. and Zafar, B. (2016). The Price is Right: Updating Inflation Expectations in a Randomized Price Information Experiment. *The Review of Economics and Statistics* 98(3): 503-523.

Bao, T., Hommes, C., Pei, J. (2021). Expectation formation in finance and macroeconomics: A review of new experimental evidence. *Journal of Behavioral and Experimental Finance* 32: 100591.

Becker, O., Leitner, J., Leopold-Wildburger, U. (2007). Heuristic modeling of expectation formation in a complex experimental information environment. *European Journal of Operational Research* 176(2): 975-985.

Bordalo, P., Gennaioli, N., Kwon, S.Y., Shleifer, A. (2021). Diagnostic bubbles. *Journal of Financial Economics* 141(3): 1060-1077.

Born, B., Enders, Z., and Muller, G.J. (2023). On FIRE, news, and expectations. Forthcoming in *The Routledge Handbook of Economic Expectations in Historical Perspective*.

Brayton, F., Mauskopf, E., Tinsley, P., Williams, J. (1997). The Role of Expectations in the FRB/US Macroeconomic Model. Federal Research Board.

Burmeister, E. and Turnovsky, S.J. (1976). The Specification of Adaptive Expectations in Continuous Time Dynamic Economic Models. *Econometrica* 44(5): 879-905.

Cavallo, A., Cruces, G., and Perez-Truglia, R. (2017). Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments. *American Economic Journal: Macroeconomics* 9(3): 1-35.

Chen, D.L., Schonger, M. and Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* 9: 88-97.

Coibion, O., Gorodnichenko, Y., Kamdar, R. (2018). The Formation of Expectations, Inflation and the Phillips Curve. *Journal of Economic Literature* 56(4): 1447-91.

Colasante, A., Palestrini, A., Russo, A. and Gallegati, M. (2017). Adaptive expectations versus rational expectations: Evidence from the lab. *International Journal of Forecasting* 33(2017): 988-1006.

Cornand, C. and Hubert, P. (2020). On the external validity of experimental inflation forecasts: A comparison with five categories of field expectations. Journal of Economic Dynamics & Control 110(2020): 103746.

Cornand, C. and Hubert, P. (2022). Information frictions across various types of inflation expectations. *European Economic Review* 146: 104175.

Czudaj, R.L. 2024. Expectation formation and the Phillips curve revisited. *Macroeconomic Dynamics*, 2024: 1-45.

Frydman, R. and Frydman, H. (2022). Why Diagnostic Expectations Cannot Replace REH. Institute of New Economic Thinking Working Paper No. 175.

Gennaioli, N. and Shleifer, A. (2018). *The Crisis of Beliefs: Investor Psychology and Financial Fragility*, Princeton University Press.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science* 15: 534–539.

Hommes, C., Sonnemans, J., Tuinstra, J., van de Velden, H. (2008). *Journal of Economic Behavior and Organization* 67(1): 116-133.

Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3(3): 430-454.

Kryvtsov, O., Petersen, L. 2021. Central bank communication that works: Lessons from lab experiments. *Journal of Monetary Economics* 117: 760-780.

L'Huillier, J-P., Singh, S.R., Yoo, D. 2023. Incorporating Diagnostic Expectations into the New Keynesian Framework. *The Review of Economic Studies,* rdad101. doi.org/10.1093/restud/rdad101.

Malmendier, U. and Nagel, S. (2015). Learning from inflation experiences. *The Quarterly Journal of Economics* 131(1): 53-87.

Mankiw, N.G. and Reis, R. (2002). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *The Quarterly Journal of Economics* 117(4): 1295–1328.

Marimon, R. and Sunder, S. (1994). Expectations and Learning under Alternative Monetary Regimes: An Experimental Approach. *Economic Theory* 4(1): 131-162.

Muth, J.F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica* 29(3): 315-335.

Rholes, R. and Petersen, L. (2021). Should central banks communicate uncertainty in their projections? *Journal of Economic Behavior and Organization* 183: 320-341.

Sugden, R., Zheng, J. and Zizzo, D.J. (2013). Not all anchors are created equal. *Journal of Economic Psychology* 39: 21-31.

Wang, T. (2024). How Do Agents Form Macroeconomic Expectations? Evidence from Inflation Uncertainty. Bank of Canada Staff Working Paper 2024-5.

Annex 1: Sample Experiment Instructions
(Low-High Volatility with Non-Binding Price Target)

Welcome to today's online experiment on decision-making.

You will earn PHP 100 for participating in today's session. You will have the opportunity to earn an additional amount of money which will depend on a series of decisions you will make and on chance. You will receive your earnings via *GCash* or *Maya* within 36 hours after the end of the experiment.

In this experiment, you will be randomly grouped with 5 other people who are also participating online. You will not have the opportunity to communicate with the members of your group.

Your group will encounter two Markets with 30 periods each. At the start of each period, your task is to forecast the market price for that period. The accuracy of your forecasts will determine how much money you will earn at the end of the experiment.

This is how the market price will be determined. For each period, the computer will use an equation. You will not know the equation, but you know that the market price per period will be calculated by the computer using information on past prices, the price forecasts of your group, and a random number.

In Market 1, that random number per period is determined based on the roll of a die with 3 sides. This means that the random number in Market 1 will range from 1 to 3. In Market 2, the random number per period is based on the roll of a die with 8 sides. For Market 2, the random number will range from 1 to 8.

For each period, you will have 30 seconds to type your price forecast on your screen. We shall use an experimental currency called ECU. You will then click on the button labelled "submit" to confirm your forecast. The minimum price is ECU 1, and the maximum price is ECU 500. At the end of each period, your screen will display a graph showing your forecast, the market price, and the target price range.

The target price range for each period will be shown as 2 diagonal lines on your screen. However, the actual market price may be higher or lower than the target price range.

In Market 1, you will have a starting fund of ECU 900. For each period, your fund will be deducted depending on the accuracy of your forecast. Refer to the table below.

If the difference between the market price and your forecast (in absolute value) is	Your ECU deduction for the period is
0	0
1 to 3	-5
4 to 6	-10
7 to 9	-15
10 to 12	-20
13 or more	-25

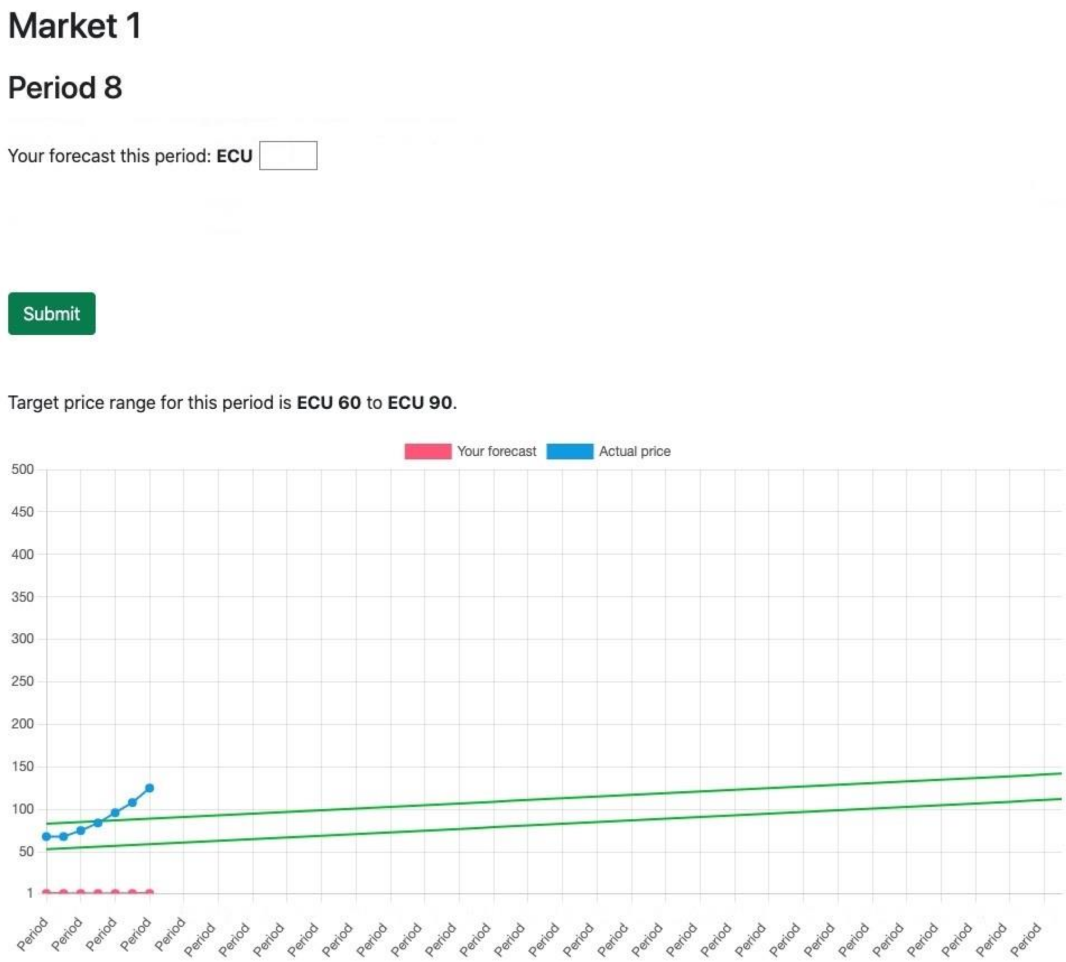
For example, in period 1 of Market 1, if the difference between your forecast and the market price is ECU 17, ECU 25 will be deducted from your fund. This means that your remaining balance at the start of period 2 will be 875. Your ending balance in period 30 will be your earnings in Market 1.

After completing all 30 periods in Market 1, the experiment will proceed with Market 2. Although your tasks in Markets 1 and 2 are similar, these two markets are not related.

Market 2 has 30 rounds. Your starting fund in Market 2 is also ECU 900. The sum of your remaining balances in Market 1 and Market 2 will be multiplied by 0.25 to determine the equivalent of your earnings in pesos. Your total earnings will be the peso equivalent, plus your show-up fee of PHP 100.

If you have a question, please raise your virtual hand in *Zoom*, and one of the experimenters will contact you to answer your question privately. If you have no question, please click on the button labelled "Next" that appears on your screen. You may refer to these instructions at any time during the experiment.

Annex 2: Sample Computer Screen
(With Non-Binding Price Target)



Annex 3: Post-Experiment Questionnaire

Thank you for participating in this experiment. Please provide all the information asked for below.

1) You consider yourself as	○ Female ○ Male ○ Non-binary
2) Your year of birth	
3) Country where you grew up	
4) Your college or university	
5) Your degree program (e.g., BS Accountancy, Juris Doctor)	
6) Before submitting your forecast in every period,	
did you consider the *actual price* in the previous period?	○ not at all ○ rarely ○ sometimes ○ often ○ all the time
did you consider *your own forecast* in the previous period?	○ not at all ○ rarely ○ sometimes ○ often ○ all the time
did you think about your *group members' forecast*?	○ not at all ○ rarely ○ sometimes ○ often ○ all the time
did you think about your accumulated earnings?	○ not at all ○ rarely ○ sometimes ○ often ○ all the time
[if appropriate] did you consider the price range of ECU 100 to 150?	○ not at all ○ rarely ○ sometimes ○ often ○ all the time
7) What other information did you consider before submitting your forecasts?	

Annex 4. Average reduction in forecast errors between markets 1 and 2

	Raw forecast errors			Absolute value of forecast errors		
	Low-High		High-Low	Low-High		High-Low
3-players	-81.76	>	-82.64	-33.73	>	-78.66
6-players	-96.54	<	-49.47	-60.47	<	-53.99
target price range	-7.52	<	-48.57	-20.14	>	-65.86

Annex 5 – Truncated autocorrelation of forecast errors by treatment

```
CONDITION: 3 players, market 1 (low volatility)
                                                 -1      0       1 1       0       1
LAG        AC       PAC       Q      Prob>Q  [Autocorrelation] [Partial Autocor]
1        0.2665   0.2673   255.83  0.0000          -- --
2        0.2419   0.1547   466.67  0.0000          -  -
3        0.1615   0.0279   560.74  0.0000          -
4        0.1429   0.0095   634.37  0.0000          -
5        0.1267  -0.0119   692.24  0.0000          -
```

```
CONDITION: 3 players, market 2 (high volatility)
                                                 -1      0       1 1       0       1
LAG        AC       PAC       Q      Prob>Q  [Autocorrelation] [Partial Autocor]
1        0.6516   0.6852   1529.7  0.0000          -----    -----
2        0.5716   0.1850     2707  0.0000          ----     -
3        0.5230   0.1384     3693  0.0000          ----     -
4        0.4946   0.1285   4575.2  0.0000          ---      -
5        0.4223  -0.0130   5218.3  0.0000          ---
```

```
CONDITION: 3 players, market 1 (high volatility)
                                                 -1      0       1 1       0       1
LAG        AC       PAC       Q      Prob>Q  [Autocorrelation] [Partial Autocor]
1        0.3218   0.3252   382.48  0.0000          -- --
2        0.2275   0.1664    573.6  0.0000          -  -
3        0.1477   0.0258   654.23  0.0000          -
4        0.1566   0.0817   744.86  0.0000          -
5        0.1287   0.0198   806.06  0.0000          -
```

```
CONDITION: 3 players, market 2 (low volatility)
                                                 -1      0       1 1       0       1
LAG        AC       PAC       Q      Prob>Q  [Autocorrelation] [Partial Autocor]
1        0.0437   0.0438   7.0588  0.0079
2        0.0089   0.0072   7.3518  0.0253
3        0.0022   0.0019   7.3699  0.0610
4       -0.0013  -0.0013   7.3759  0.1173
5        0.0007   0.0011   7.3776  0.1940
```

```
CONDITION: 6 players, market 1 (low volatility)
                                                 -1      0       1 1       0       1
LAG        AC       PAC       Q      Prob>Q  [Autocorrelation] [Partial Autocor]
1        0.3114   0.3124   349.27  0.0000          -- --
2        0.1838   0.0423   470.98  0.0000          -
3        0.1237  -0.0023   526.12  0.0000
4        0.1032  -0.0027   564.54  0.0000
5        0.0951   0.0052   597.18  0.0000
```

CONDITION: 6 players, market 2 (high volatility)

					-1 0 1	1 0 1
LAG	AC	PAC	Q	Prob>Q	[Autocorrelation]	[Partial Autocor]
1	0.0353	0.0354	4.4775	0.0343		
2	-0.0069	-0.0074	4.6496	0.0978		
3	-0.0035	-0.0023	4.6937	0.1957		
4	-0.0097	-0.0095	5.0342	0.2838		
5	-0.0027	-0.0019	5.0603	0.4086		

CONDITION: 6 players, market 1 (high volatility)

					-1 0 1	1 0 1
LAG	AC	PAC	Q	Prob>Q	[Autocorrelation]	[Partial Autocor]
1	0.4815	0.4876	835.31	0.0000	---	---
2	0.3713	0.0553	1332.2	0.0000	--	
3	0.3711	0.0906	1828.7	0.0000	--	
4	0.3477	-0.0190	2264.6	0.0000	--	
5	0.3020	-0.0228	2593.6	0.0000	--	

CONDITION: 6 players, market 2 (low volatility)

					-1 0 1	1 0 1
LAG	AC	PAC	Q	Prob>Q	[Autocorrelation]	[Partial Autocor]
1	0.7206	0.7219	1871.1	0.0000	-----	-----
2	0.6705	0.3035	3491.4	0.0000	-----	--
3	0.6413	0.1669	4974	0.0000	-----	-
4	0.6565	0.2719	6528.4	0.0000	-----	--
5	0.6158	0.1609	7896	0.0000	----	-

CONDITION: 6 players with target price range, market 1 (low volatility)

					-1 0 1	1 0 1
LAG	AC	PAC	Q	Prob>Q	[Autocorrelation]	[Partial Autocor]
1	0.0770	0.0772	22.422	0.0000		
2	0.0241	0.0171	24.622	0.0000		
3	0.0725	0.0704	44.515	0.0000		
4	0.0043	-0.0215	44.584	0.0000		
5	-0.0036	-0.0049	44.633	0.0000		

CONDITION: 6 players with target price range, market 2 (high volatility)

					-1 0 1	1 0 1
LAG	AC	PAC	Q	Prob>Q	[Autocorrelation]	[Partial Autocor]
1	0.1306	0.1393	64.544	0.0000	- -	
2	0.1684	0.1960	171.8	0.0000	- -	
3	0.0912	0.0617	203.28	0.0000		
4	0.0805	0.0430	227.8	0.0000		
5	0.0720	0.0514	247.46	0.0000		

CONDITION: 6 players with target price range, market 1 (high volatility)

					-1 0 1	1 0 1
LAG	AC	PAC	Q	Prob>Q	[Autocorrelation]	[Partial Autocor]
1	0.2900	0.2925	303.05	0.0000	-- --	
2	0.1357	0.0469	369.43	0.0000	-	
3	0.0257	-0.0081	371.82	0.0000		
4	0.0219	-0.0032	373.56	0.0000		
5	0.0076	-0.0139	373.76	0.0000		

CONDITION: 6 players with target price range, market 2 (high volatility)

					-1 0 1	1 0 1
LAG	AC	PAC	Q	Prob>Q	[Autocorrelation]	[Partial Autocor]
1	0.0211	0.0211	1.6055	0.2051		
2	0.0057	0.0056	1.7218	0.4228		
3	-0.0159	-0.0169	2.6383	0.4508		
4	0.0093	0.0100	2.9502	0.5662		
5	0.0008	0.0010	2.9523	0.7073		