On-line Appendix for Asset price volatility and investment horizons: An experimental investigation

Mikhail Anufriev Aleksei Chernulich Jan Tuinstra

8 January 2021

In this Appendix we repeat illustrations and tables from the paper, but include all the data that we collected in the experimental sessions. In contract to the main text, we do not exclude those outliers that we decided not to report in the paper.

Outliers

As explained in Sections 2.1 and 3.1 (see in particular footnote 21) of the paper, the LtF experiments are vulnerable to 'outliers', i.e., abnormal behavior of one participant that strongly affect the price dynamics in the whole group. The outliers can be either stand alone individual forecasts laying very far from the previous prices,[1] or whole patterns of forecasting behavior of a participant that do not seem to be connected to the past price dynamics. Authors often disregard the periods or whole groups affected by such behavior.

In our paper we had to do the same. Indeed, we decided to consider as outliers (i) one group (group 4) in the **ConH1** treatment, (ii) one group (group 6) in the **OscH2** treatment, and (iii) the 10 last periods in group 1 of the **OscH3SF** treatment.

Price and predictions in the former two groups are shown in Fig. A1. In group 4 of the **ConH1** treatment, participant 2 during the whole experiment

[1] There could be several reasons why a participant may submit such a forecast, including motives to investigate the reaction of the system to the extreme forecast, but also a simple typo. The forecast outliers usually lead to lower rewards for all participants in the group, though the participant who submitted the outlier forecast typically suffers most.

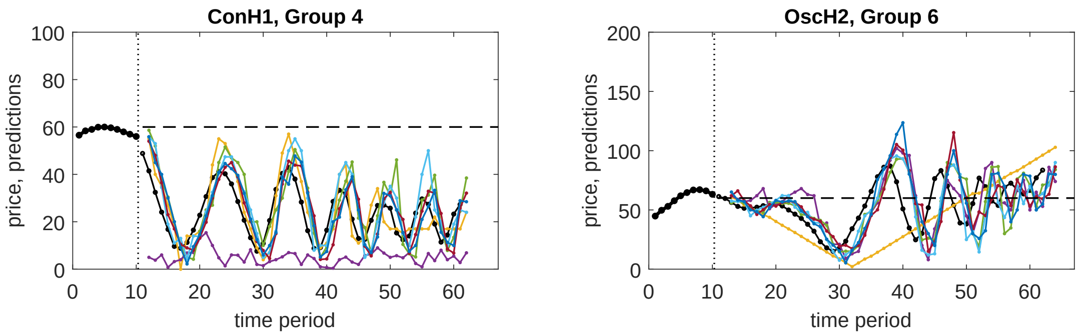

Figure A1: Group outliers of the experiment.

predicted an extremely low price, i.e., 5.0 on average and not higher than 15.4. In group 6 of the **OscH2** treatment, participant 1 during the first 20 decision periods predicted a price monotonically decreasing from 60.59 to 2.15 and during the remaining periods predicted a constantly increasing price up to 102.85. In both cases the participants' behaviors were disconnected from the past prices, and affected the outcome. In particular, prices dropped significantly in the beginning of the experiment in both groups in patterns that were not observed in other groups.

The dynamics in group 1 of the **OscH3SF** treatment in shown in Fig. 17, Appendix D of the paper. In that group, participant 5 towards the end of the experiment started to submit extreme predictions, i.e., predictions equal to either 0, 999 or 1000. These extreme predictions have a large impact on the realized market price in periods $54-63$ in that group.

Descriptive statistics and illustrations

Table A1 collects various descriptive statistics of the experimental data. The statistics are computed only for those 50 periods for which participants are paid for the accuracy of their forecasts. The precise periods for each treatment are given in the penultimate column of Table 1 of the paper.

We show the following statistics (in the order of columns) for each experimental group: mean price, standard deviation of prices, the median RAD, i.e., the median (over all incentivized periods) of the relative absolute deviation of price from the fundamental price, $|p_t - p^f|/p^f$, the IQR (interquartile range), i.e., the length of the interval that contains the middle 50% of the prices, the discoordination measure defined as the median of the standard deviations of predictions, and the average payoff per participant in points, as defined in Eq. (2) of the paper. The average

Treatment	Group	Mean price	Std dev	Median RAD	IQR	Discoord	Average payoffs
ConH1	Group 1	57.67	2.07	0.04	3.36	0.66	1221.2
	Group 2	58.69	0.99	0.02	1.23	0.28	1272.4
	Group 3	58.23	3.65	0.02	3.72	0.46	1226.5
	Group 4	23.29	10.33	0.62	16.64	10.00	301.4
	Average	49.47	4.26	0.17	6.24	2.85	1005.4
	Median	57.95	2.86	0.03	3.54	0.56	1223.8
ConH2	Group 1	56.63	0.99	0.05	1.35	1.01	1198.9
	Group 2	64.33	14.34	0.04	3.99	1.97	857.6
	Group 3	59.37	1.08	0.01	0.81	0.70	1215.5
	Group 4	59.08	0.70	0.01	0.88	0.68	1241.1
	Average	59.85	4.28	0.03	1.76	1.09	1128.3
	Median	59.22	1.04	0.03	1.12	0.86	1207.2
ConH3	Group 1	63.46	1.28	0.06	1.47	2.12	1016.7
	Group 2	58.96	1.24	0.02	1.82	1.31	1121.6
	Group 3	59.76	0.62	0.00	0.80	0.62	1252.2
	Average	60.73	1.05	0.03	1.36	1.35	1130.2
	Median	59.76	1.24	0.02	1.47	1.31	1121.6
OscH1	Group 1	63.28	1.26	0.05	0.86	0.37	1266.0
	Group 2	60.64	1.64	0.02	2.76	0.88	1223.5
	Group 3	52.25	23.91	0.36	42.42	5.24	494.7
	Group 4	63.85	6.96	0.08	6.48	1.39	946.3
	Group 5	61.30	1.07	0.02	0.99	0.35	1266.9
	Group 6	72.25	45.16	0.49	76.90	5.44	382.9
	Group 7	231.63	305.74	0.47	268.87	6.18	475.5
	Group 8	67.64	57.04	0.11	16.78	0.87	909.5
	Average	84.10	55.35	0.20	52.01	2.59	870.7
	Median	63.57	15.43	0.10	11.63	1.14	927.9
OscH2	Group 1	54.84	6.76	0.11	8.98	4.94	413.5
	Group 2	51.93	18.56	0.16	23.67	5.27	332.3
	Group 3	58.42	14.08	0.14	15.99	6.29	279.0
	Group 4	60.32	11.83	0.19	23.46	5.67	213.5
	Group 5	57.84	13.09	0.16	19.86	9.79	149.9
	Group 6	53.51	19.04	0.22	30.16	11.93	246.2
	Average	56.14	13.89	0.16	20.35	7.32	272.4
	Median	56.34	13.58	0.16	21.66	5.98	262.6
OscH3	Group 1	60.47	2.18	0.02	2.40	2.50	977.6
	Group 2	67.57	12.73	0.09	2.49	2.76	791.4
	Group 3	60.56	2.57	0.03	3.66	1.77	813.4
	Group 4	54.37	11.12	0.12	10.90	10.31	270.6
	Group 5	61.33	4.60	0.03	2.28	3.89	766.6
	Group 6	58.57	3.61	0.02	2.20	1.53	907.9
	Average	60.48	6.14	0.05	3.99	3.79	754.6
	Median	60.52	4.11	0.03	2.44	2.63	802.4
OscH3SF	Group 1	88.92	63.48	0.02	6.78	0.76	778.9
	Group 2	60.72	0.97	0.01	1.02	0.88	1192.8
	Group 3	59.76	15.61	0.14	17.11	7.21	284.2
	Group 4	55.73	1.82	0.07	2.24	2.21	1022.2
	Group 5	50.60	19.54	0.19	21.05	11.96	206.4
	Group 6	45.10	14.72	0.23	17.31	8.61	249.0
	Average	60.14	19.35	0.11	10.92	5.27	622.2
	Median	57.74	15.16	0.11	11.95	4.71	531.5

Table A1: Descriptive statistics in the experiment. No outliers removed.

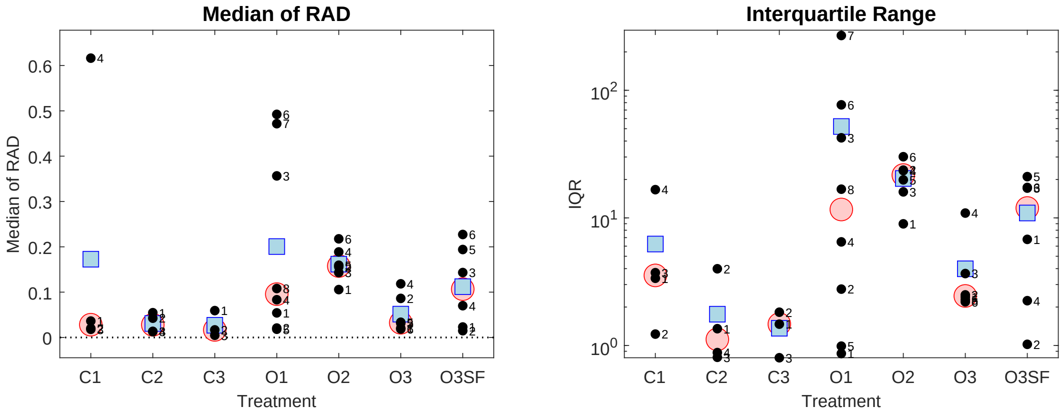

Figure A2: *Left:* Level of mispricing measured as the median of the relative absolute deviation. *Right:* Level of volatility measured as the Interquartile Range (log scale). No outliers removed.

and median of statistics for all groups within each treatment are also reported in Table A1. Overall, this is the same table as Table 2 from Appendix C.1 of the paper, except that outliers are not removed.

In the paper we illustrate the descriptive statistics in Fig. 5, where the left panel shows the median of RAD and the right panel shows the IQR. We replicate these illustrations using all data in Fig. A2. In both panels, statistics for groups are organized by treatments, where the statistics for individual groups are indicated by the black dots. The average and the median for each treatment are shown by the blue square or red circle, respectively.

Nonparametric tests for statistics comparisons

The results in the paper are supported by the Kolmogorov-Smirnov, Fischer-Pitman and Mann-Whitney-Wilcoxon tests that compare the measures pairwise between the treatments. In particular, the p-values of these tests for the median of RAD and the IQR comparisons are reported in Appendices C.2 and C.3, respectively.

We run the same tests with the complete data, and report the p-values in Tables A2-A4 for the median of RAD, and in Tables A5-A7 for the IQR.[2] As in the paper, *, **, and *** in these tables denote treatment comparisons when the

[2]The pairwise comparisons for treatments **ConH2**, **ConH3**, **OscH1** and **OscH3** are not affected by the outliers. For those comparisons, the p-values coincide with the p-values for the corresponding tests in the main paper.

null hypothesis of inequality is rejected at the 10%, 5% or 1% significance level, respectively.

Note that heterogeneity in the statistics within the same treatment may prevent us from getting significant results. Adding outliers increases heterogeneity within the treatment and thus typically reduces the number of significant pairwise comparisons. This is especially the case for comparisons that involve the **ConH1** treatment. Indeed, group 4 in this treatment was identified as outlier due to a behavior of one of the participants. Fig. A2 shows that the values of both statistics for this group are also far from the corresponding statistic in the other groups of the same treatment. Thus, comparisons involving this treatment will likely be insignificant. On the other hand, the results in the paper are not strongly based on significant comparisons involving the **ConH1** treatment, whereas other outliers do not have a similar effect for other treatments. Generally, inspecting the tests based on all data, we observe that the major tendencies and results remain even after including all outliers.

Result 1 of the paper, about the effect of initial price history, for instance, is strongly based upon a significant differences for the median RAD and IQR comparisons between treatments **ConH2** and **OscH2**. These statistics are significantly different at the 1% level for all three tests even when outliers are not removed. Similarly, Result 2 about the effect of investment horizon, is strongly based upon a significant differences when the median RAD and IQR are compared for treatments **OscH2** and **OscH3**. All significant differences for the three tests (at least at the 5% levels) in these comparisons remains.

Median RAD	**ConH1**	**ConH2**	**ConH3**	**OscH1**	**OscH2**	**OscH3**	**OscH3SF**
ConH1	X	0.723	0.969	0.191	0.033**	0.510	0.349
ConH2	0.273	X	0.601	0.075*	0.002***	0.220	0.068*
ConH3	0.131	0.601	X	0.083*	0.006***	0.099*	0.099*
OscH1	0.661	1.000	1.000	X	0.127	0.879	0.879
OscH2	0.685	1.000	1.000	0.313	X	1.000	0.816
OscH3	0.685	0.845	1.000	0.239	0.006***	X	0.160
OscH3SF	0.685	1.000	1.000	0.313	0.160	0.816	X

Table A2: p-values of the one-sided Kolmogorov-Smirnov test comparing the median of RAD statistics, pairwise for different treatments, where alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers are removed.

Median RAD	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.676	0.694	0.474	0.588	0.682	0.635
ConH2	0.324	X	0.556	0.030**	0.005***	0.227	0.066*
ConH3	0.306	0.444	X	0.048**	0.012**	0.212	0.071*
OscH1	0.526	0.970	0.952	X	0.676	0.936	0.832
OscH2	0.412	0.995	0.988	0.324	X	0.998	0.883
OscH3	0.318	0.773	0.788	0.064*	0.002***	X	0.085*
OscH3SF	0.365	0.934	0.929	0.168	0.117	0.915	X

Table A3: p-values of the one-sided Fischer-Pitman permutation test comparing the median of RAD statistics, pairwise for different treatments, where alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers are removed.

Median RAD	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.757	0.886	0.230	0.129	0.457	0.381
ConH2	0.343	X	0.571	0.036**	0.005***	0.238	0.057*
ConH3	0.200	0.571	X	0.042**	0.012**	0.131	0.083*
OscH1	0.816	0.976	0.976	X	0.286	0.886	0.714
OscH2	0.914	1.000	1.000	0.755	X	0.999	0.803
OscH3	0.619	0.824	0.917	0.141	0.002***	X	0.197
OscH3SF	0.695	0.967	0.952	0.331	0.242	0.845	X

Table A4: p-values of the one-sided Mann-Whitney-Wilcoxon test comparing the median of RAD statistics, pairwise for different treatments, where alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers are removed.

IQR	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	1.000	1.000	0.191	0.033**	0.685	0.220
ConH2	0.273	X	0.451	0.075*	0.002***	0.033**	0.068*
ConH3	0.076*	0.601	X	0.043**	0.006***	0.006***	0.027**
OscH1	0.661	1.000	1.000	X	0.127	0.597	0.597
OscH2	1.000	1.000	1.000	0.313	X	1.000	1.000
OscH3	0.349	0.959	1.000	0.127	0.006***	X	0.160
OscH3SF	0.845	1.000	1.000	0.313	0.160	0.816	X

Table A5: p-values of the one-sided Kolmogorov-Smirnov test comparing the interquartile range statistics, pairwise for different treatments, where alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers are removed.

Estimated AR(2) Rules

In the main paper we fitted the AR(2) linear rule

$$p_t = \beta_0 + \beta_1 p_{t-1} + \beta_2 p_{t-2} + \nu_t \,,$$

IQR	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.873	0.917	0.145	0.014**	0.754	0.180
ConH2	0.127	X	0.528	0.042**	0.005***	0.104	0.033**
ConH3	0.083*	0.472	X	0.048**	0.012**	0.012**	0.047**
OscH1	0.855	0.958	0.952	X	0.718	0.953	0.855
OscH2	0.986	0.995	0.988	0.282	X	0.998	0.966
OscH3	0.246	0.896	0.988	0.047**	0.002***	X	0.052*
OscH3SF	0.820	0.967	0.953	0.145	0.034**	0.948	X

Table A6: p-values of the one-sided Fischer-Pitman permutation test comparing the interquartile range statistics, pairwise for different treatments, where alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers are removed.

IQR	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.900	0.943	0.285	0.019**	0.762	0.238
ConH2	0.171	X	0.571	0.055*	0.005***	0.086*	0.033**
ConH3	0.114	0.571	X	0.067*	0.012**	0.012**	0.048**
OscH1	0.770	0.964	0.958	X	0.331	0.886	0.574
OscH2	0.990	1.000	1.000	0.714	X	0.999	0.968
OscH3	0.305	0.943	1.000	0.141	0.002***	X	0.197
OscH3SF	0.824	0.981	0.976	0.475	0.047**	0.845	X

Table A7: p-values of the one-sided Mann-Whitney-Wilcoxon test comparing the interquartile range statistics, pairwise for different treatments, where alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers are removed.

see Eq. (5) in the paper, for the price dynamics in each group. The results of estimations are collected in Table 9 in Appendix C.4 and illustrated in the stability triangles in Fig. 6.

Table A8 collects estimations of the AR(2) rules when no outliers are removed. It is different from Table 9 of the paper only because it has two more experimental groups and uses all available periods to estimate the rule in group 1 of treatmet **OscH3SF**. As in the paper we denote by *, **, and *** the coefficients that are significant at the 10%, 5% or 1% significance level, respectively. The standard errors are shown below in parentheses. Two columns called 'autoc' and 'hskd' collect the p-values of the specification tests for the residual autocorrelation (Ljung-Box test) and heteroscedasticity (Engle's ARCH tests), respectively. The group label is shown in bold when the estimated rule passes both tests at the 5% significance level. In the main text we notice that the price dynamics in 26 of the 35 markets can be described well by the AR(2) rule, in the sense that the estimated rules pass two specification tests (on autocorrelation and heteroskedasticity) for residuals. Using the whole data set we find that 27 of the 37 markets are described by the AR(2) rule well.

The estimations are illustrated using the stability triangles in Fig. A3. We show only the treatments that are affected by the outliers, as for all other treatments there are no changes with respect to Fig. 6 of the paper.

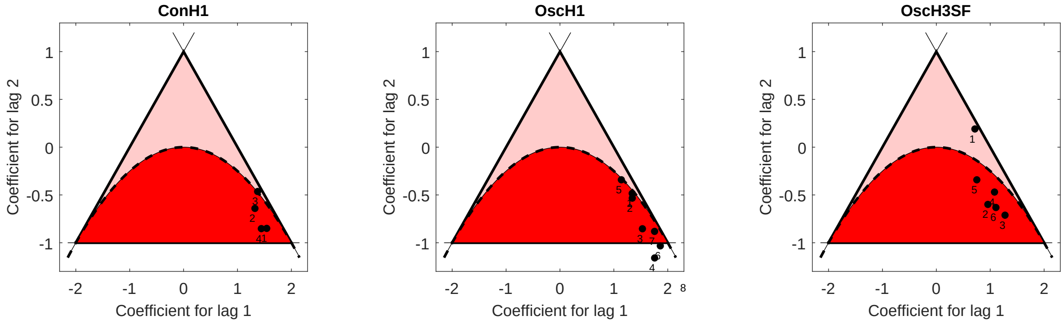

Figure A3: Stability triangles for AR(2) model, Eq. (5) of the paper, and estimated AR(2) rules with no outliers removed. Three treatments that are affected by the outliers are shown.

Treatment	Group	Const	Past Prices		Specification Tests		AdjRSq	NObs
			lag 1	lag 2	autoc	hskd		
ConH1	**Group 1**	17.72***	1.54***	−0.85***	0.71	0.35	0.92	48
		(2.55)	(0.08)	(0.08)				
	Group 2	18.66***	1.32***	−0.64***	0.68	0.70	0.78	48
		(4.24)	(0.11)	(0.11)				
	Group 3	5.41***	1.37***	−0.46***	0.34	0.24	0.97	48
		(1.39)	(0.12)	(0.11)				
	Group 4	9.58***	1.44***	−0.85***	0.20	0.23	0.90	48
		(1.17)	(0.07)	(0.07)				
ConH2	**Group 1**	40.99***	0.90***	−0.62***	0.13	0.25	0.57	48
		(6.05)	(0.12)	(0.11)				
	Group 2	45.76***	0.44***	−0.15	1.00	0.88	0.13	48
		(10.70)	(0.15)	(0.15)				
	Group 3	53.60***	0.49***	−0.39***	0.99	0.54	0.23	48
		(9.24)	(0.14)	(0.14)				
	Group 4	52.03***	0.54***	−0.42***	0.74	0.43	0.28	48
		(8.32)	(0.13)	(0.13)				
ConH3	**Group 1**	35.45***	0.91***	−0.47***	0.51	0.57	0.51	48
		(7.24)	(0.13)	(0.13)				
	Group 2	27.81***	1.21***	−0.69***	0.69	0.54	0.73	48
		(4.80)	(0.11)	(0.11)				
	Group 3	46.64***	0.48***	−0.26*	0.32	0.07	0.17	48
		(9.50)	(0.14)	(0.14)				
OscH1	**Group 1**	9.05***	1.34***	−0.48***	0.42	0.88	0.87	48
		(3.13)	(0.13)	(0.12)				
	Group 2	11.66***	1.34***	−0.53***	0.81	0.39	0.84	48
		(3.71)	(0.12)	(0.12)				
	Group 3	16.87***	1.53***	−0.85***	0.01	0.31	0.91	48
		(2.68)	(0.08)	(0.08)				
	Group 4	25.34***	1.76***	−1.16***	0.06	0.85	0.98	48
		(1.29)	(0.04)	(0.04)				
	Group 5	12.36***	1.14***	−0.34**	0.66	0.52	0.76	48
		(4.19)	(0.14)	(0.13)				
	Group 6	12.29***	1.86***	−1.03***	0.00	0.44	0.99	48
		(1.48)	(0.04)	(0.04)				
	Group 7	29.91**	1.76***	−0.88***	0.08	0.06	0.96	48
		(11.32)	(0.08)	(0.09)				
	Group 8	2.56**	2.33***	−1.37***	0.92	0.55	1.00	48
		(1.13)	(0.04)	(0.05)				
OscH2	**Group 1**	28.23***	1.28***	−0.79***	0.74	0.23	0.83	48
		(3.43)	(0.08)	(0.08)				
	Group 2	25.82***	1.40***	−0.89***	0.03	0.55	0.89	48
		(2.85)	(0.07)	(0.07)				
	Group 3	28.71***	1.31***	−0.80***	0.01	0.00	0.82	48
		(3.95)	(0.09)	(0.09)				
	Group 4	34.97***	1.28***	−0.86***	0.09	0.02	0.86	48
		(3.56)	(0.07)	(0.07)				
	Group 5	33.02***	1.19***	−0.76***	0.03	0.84	0.76	48
		(4.65)	(0.10)	(0.10)				
	Group 6	13.89***	1.51***	−0.76***	0.32	0.63	0.87	48
		(3.15)	(0.10)	(0.10)				
OscH3	**Group 1**	23.33***	0.59***	0.02	0.97	0.67	0.37	48
		(7.12)	(0.15)	(0.14)				
	Group 2	47.17***	0.40***	−0.10	1.00	0.88	0.10	48
		(11.37)	(0.15)	(0.15)				
	Group 3	21.56***	1.53***	−0.89***	0.12	0.41	0.92	48
		(2.64)	(0.07)	(0.07)				
	Group 4	34.48***	0.94***	−0.56***	0.16	0.19	0.60	48
		(4.90)	(0.11)	(0.10)				
	Group 5	39.42***	0.99***	−0.63***	0.00	0.33	0.69	48
		(4.63)	(0.10)	(0.08)				
	Group 6	10.59***	1.54***	−0.71***	0.01	0.10	0.93	48
		(2.09)	(0.09)	(0.09)				
OscH3SF	Group 1	12.08	0.72***	0.19	0.01	0.01	0.67	48
		(9.81)	(0.15)	(0.17)				
	Group 2	38.99***	0.96***	−0.60***	0.21	0.93	0.61	48
		(5.56)	(0.11)	(0.10)				
	Group 3	26.25***	1.27***	−0.71***	0.46	0.13	0.77	48
		(4.61)	(0.10)	(0.10)				
	Group 4	21.73***	1.08***	−0.47***	0.13	0.60	0.66	48
		(5.12)	(0.12)	(0.11)				
	Group 5	30.25***	0.75***	−0.34**	1.00	0.98	0.36	48
		(7.03)	(0.14)	(0.14)				
	Group 6	23.44***	1.11***	−0.63***	0.26	0.05	0.66	48
		(4.29)	(0.11)	(0.11)				

Table A8: Estimated AR(2) rules. No outliers removed.

Trend-extrapolation

In Section 3.2 of the paper we investigate how the degree of trend extrapolation changes with the investment horizons. For this purpose we draw in Fig. (8) the scatter plots, where the averages of participants' expected price changes, $\bar{p}^e_{t+H} - p_{t-1}$, are plotted against the last observed price change, $p_{t-1} - p_{t-2}$.

Fig. A4 shows the same plots for those three treatments that are affected by the outliers. As before, the red line in each panel represents a linear regression line, the slope of which is shown in the title of the panel. We observe that when all data are used, the slope of the regression line decreases from 0.84 to 0.72 in the **ConH1** treatment, slightly increases from 0.64 to 0.65 in the **OscH2** treatment, and decreases for, 0.24 to -0.07 in the **OscH3SF** treatment. Importantly, the tendency of the extrapolation strength to decrease with the horizon, reported in the paper, remains. Table A9 reports the results of the t-test for the whole data, mirroring Table 16 from Appendix C.7 of the paper. It shows that the difference in slope for treatment **OscH1** with the slopes in treatments **OscH2**, **OscH3** and **OscH3SF** is more significant, when outliers are not excluded, confirming Result 3.

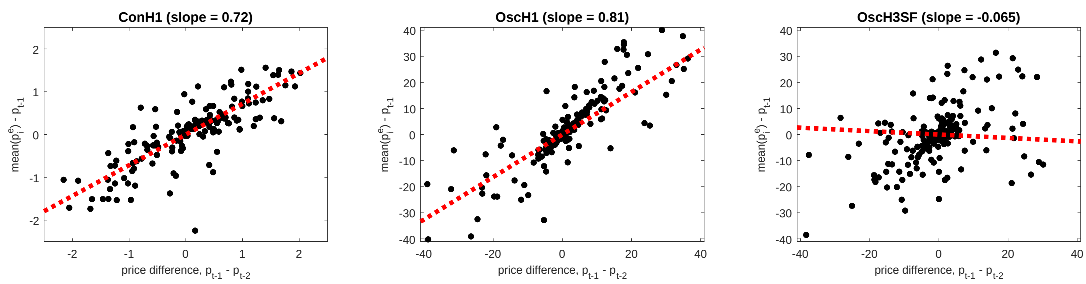

Figure A4: Scatter plots and correlations between $\bar{p}^e_{i,t+h} - p_{t-1}$ and the most recent price change, $p_{t-1} - p_{t-2}$. No outliers removed. Only the three groups that are affected by the outliers are shown.

Slope	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.179	0.002***	0.887	0.186	0.174	0.842
ConH2	0.963	X	0.963	0.988	0.973	0.970	0.985
ConH3	0.170	0.866	X	0.964	0.885	0.868	0.944
OscH1	0.031**	0.033**	0.036**	X	0.024**	0.024**	0.033**
OscH2	0.830	0.884	0.834	0.938	X	0.876	0.938
OscH3	0.182	0.216	0.169	0.926	0.785	X	0.831
OscH3SF	0.216	0.218	0.219	0.991	0.771	0.206	X

Table A9: p-values of the t-test of the differences in correlations between $\bar{p}^e_{i,t+h} - p_{t-1}$ and the most recent price change, $p_{t-1} - p_{t-2}$, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers removed.

Discoordination

In Section 3.2 of the paper, we introduced the *discoordination* measure, i.e., the median (over 50 incentivized periods) of the standard deviations of the six predictions for each period. The measure was illustrated in Fig. 9. This measure is similarly illustrated in Fig. A5 using all data. Similarly to Fig. A2 above, the diagram is organized by treatments, where the statistics for individual groups are indicated by the black dots. The average and the median for each treatment are shown by the blue square or red circle, respectively.

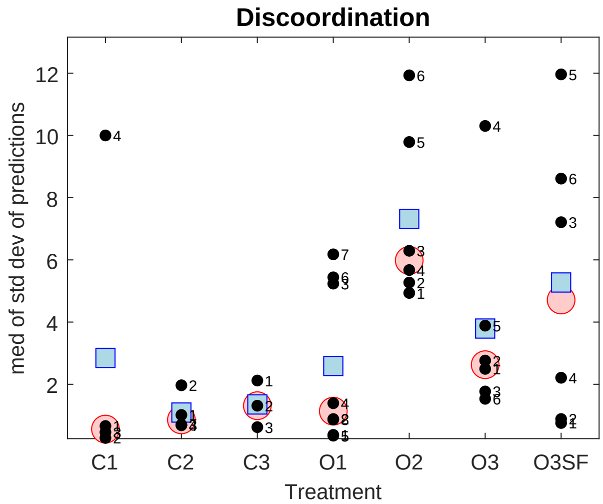

Figure A5: Level of discoordination of predictions in all markets measured as the median of the standard deviation of predictions. No outliers removed.

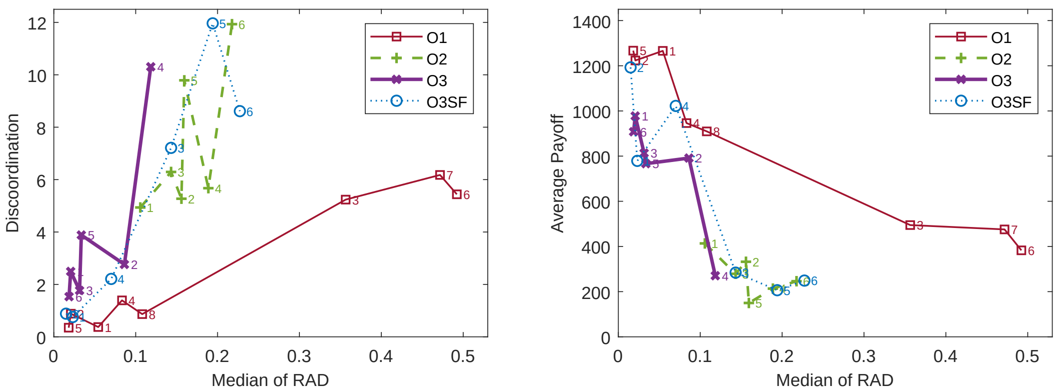

Figure A6: *Left:* Median of the standard deviation of predictions plotted against median of RAD. *Right:* Average payoff plotted against median of RAD. The data from the **Osc** treatments are used. No outliers removed.

The paper illustrates the relationship between discoordination and price volatility, as well as between discoordination and average payoffs in Fig. 9. We show the corresponding plots in Fig. A6, using all data from the **Osc** treatments. These scatter plots suggest that, controlling for the level of volatility, first, coordination gets easier, and, second, the earnings get higher, when the investment horizon is shorter. These tendencies, reported in the main paper, are not affected when we extend the scatter plots by including the outliers.

Result 4 of the paper, about the coordination of individual expectations, is also supported by the results of the three non-parametric tests that compare the discoordination measure pairwise between the treatments. The p-values of the tests are reported in Appendix C.5 of the paper. Tables A10-A12 give the test results when all data are used. As usually, *, **, and *** indicate the treatment pairs for which the null hypothesis of inequality of the discoordination is rejected at the 10%, 5% or 1% significance level, respectively. Focusing only on the **OscH1**, **OscH2** and **OscH3** treatments, we observe that there are even more significant differences, supporting Result 4, than in the main paper.

Discoord	**ConH1**	**ConH2**	**ConH3**	**OscH1**	**OscH2**	**OscH3**	**OscH3SF**
ConH1	X	0.054*	0.318	0.191	0.033**	0.033**	0.033**
ConH2	0.723	X	0.451	0.394	0.002***	0.033**	0.068*
ConH3	0.751	0.601	X	0.456	0.006***	0.099*	0.099*
OscH1	0.661	0.661	0.705	X	0.040**	0.040**	0.127
OscH2	0.959	1.000	1.000	1.000	X	1.000	0.816
OscH3	0.959	1.000	1.000	0.699	0.006***	X	0.442
OscH3SF	0.959	1.000	1.000	1.000	0.160	0.442	X

Table A10: p-values of the one-sided Kolmogorov-Smirnov test comparing the discoordination (median of the standard deviations of predictions) statistics, pairwise for different treatments, where alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers are removed.

Discoord	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.507	0.472	0.500	0.047**	0.336	0.218
ConH2	0.493	X	0.278	0.200	0.005***	0.014**	0.052*
ConH3	0.528	0.722	X	0.301	0.012**	0.047**	0.129
OscH1	0.500	0.800	0.699	X	0.007***	0.236	0.099*
OscH2	0.953	0.995	0.988	0.993	X	0.956	0.813
OscH3	0.664	0.986	0.953	0.764	0.044**	X	0.266
OscH3SF	0.782	0.948	0.871	0.901	0.187	0.734	X

Table A11: p-values of the one-sided Fischer-Pitman permutation test comparing the discoordination (median of the standard deviations of predictions) statistics, pairwise for different treatments, where alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers are removed.

Discoord	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.171	0.314	0.285	0.086*	0.086*	0.086*
ConH2	0.900	X	0.429	0.341	0.005***	0.019**	0.057*
ConH3	0.800	0.686	X	0.461	0.012**	0.048**	0.131
OscH1	0.770	0.715	0.612	X	0.010***	0.141	0.091*
OscH2	0.943	1.000	1.000	0.994	X	0.987	0.758
OscH3	0.943	0.990	0.976	0.886	0.021**	X	0.531
OscH3SF	0.943	0.967	0.917	0.929	0.294	0.531	X

Table A12: p-values of the one-sided Mann-Whitney-Wilcoxon test comparing the discoordination (median of the standard deviations of predictions) statistics, pairwise for different treatments, where alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively. No outliers are removed.