For Online Publication

A1 Proofs

First, we formally establish our earlier contention that $\hat{k}$ is independent of level-0 beliefs and of the value of the constant γ.

Lemma A.1. *Fix β and weight system ω. Let $\hat{k}(\beta,\omega,a,\gamma)$ be the optimal sophistication level given the constant term γ and level-0 beliefs a. Then $\hat{k}(\beta,\omega,a,\gamma) = \hat{k}(\beta,\omega,1,0)$.*

Proof. Write $\mathcal{T}_\gamma(a,\omega,\beta)$ as the realized value of y given the datum (β,ω,a,γ). From equation (2) we have

$$\begin{aligned}\mathcal{T}_\gamma(a,\omega,\beta) - \frac{\gamma}{1-\beta} &= \gamma + \frac{\beta\gamma}{1-\beta}\sum_{k\geq 0}\omega_k - \frac{\beta\gamma}{1-\beta}\sum_{k\geq 0}\beta^k\omega_k + \beta a\sum_{k\geq 0}\beta^k\omega_k - \frac{\gamma}{1-\beta}\\ &= \frac{\gamma}{1-\beta} - \left(\frac{\gamma}{1-\beta}\right)\beta\sum_{k\geq 0}\beta^k\omega_k + \beta a\sum_{k\geq 0}\beta^k\omega_k - \frac{\gamma}{1-\beta}\\ &= \mathcal{T}_0\left(a - \frac{\gamma}{1-\beta},\omega,\beta\right).\end{aligned} \tag{A1}$$

Next, let $\phi(\beta,k,a,\gamma)$ be the forecast of a k-level agent. Then

$$\phi(\beta,k,a,\gamma) = \gamma\left(\frac{1-\beta^k}{1-\beta}\right) + \beta^k a.$$

Also, let $\phi^\varepsilon(\beta,k,a,\gamma) = |\phi(\beta,k,a,\gamma) - \mathcal{T}_\gamma(a,\omega,\beta)|$ be the associated forecast error.

Now observe that

$$\begin{aligned}\arg\min_{k\in\mathbb{N}}\phi^\varepsilon(\beta,k,a,0) &= \arg\min_{k\in\mathbb{N}}\left|a\beta^k - a\beta\sum_{n\geq 0}\beta^n\omega_n\right| = \arg\min_{k\in\mathbb{N}}|a|\left|\beta^k - \beta\sum_{n\geq 0}\beta^n\omega_n\right|\\ &= \arg\min_{k\in\mathbb{N}}\left|\beta^k - \beta\sum_{n>0}\beta^n\omega_n\right| = \arg\min_{k\in\mathbb{N}}\phi^\varepsilon(\beta,k,1,0).\end{aligned} \tag{A2}$$

Also, by (A1) we have that

$$\phi^\varepsilon(\beta,k,a,\gamma) = \phi^\varepsilon(\beta,k,a-\bar{y},0),$$

where $\bar{y} = \gamma(1-\beta)^{-1}$, so that

$$\arg\min_{k\in\mathbb{N}}\phi^\varepsilon(\beta,k,a,\gamma) = \arg\min_{k\in\mathbb{N}}\phi^\varepsilon(\beta,k,a-\bar{y},0). \tag{A3}$$

Putting (A2) and (A3) together yields

$$\arg\min_{k\in\mathbb{N}}\phi^\varepsilon(\beta,k,a,\gamma) = \arg\min_{k\in\mathbb{N}}\phi^\varepsilon(\beta,k,1,0),$$

which completes the proof. ■

Stability of unified dynamics. The strategy is to show that adaptive dynamics lead to convergence for any sequence of weights. Some notation is needed. Given a system of weights $\omega = \{\omega_i\}_{i\geq 0}$, let

$$\mathcal{T}_\gamma(a,\omega,\beta) = \gamma\left(1+\frac{\beta}{1-\beta}\sum_{k\geq 0}(1-\beta^k)\omega_k\right) + \beta\sum_{k\geq 0}\beta^k\omega_k a \tag{A4}$$

Now fix any *sequence* of weight systems $\{\omega_t\}_{t\geq 0} = \left\{\{\omega_{it}\}_{i\geq 0}\right\}_{t\geq 0}$, and define the following recursion:

$$a_t = a_{t-1} + \phi\left(\mathcal{T}_\gamma\left(a_{t-1},\omega_{t-1},\beta\right) - a_{t-1}\right). \tag{A5}$$

We have the following result.

Lemma A.2. *Let $\phi \in (0,1]$.*

1. *If $|\beta| < 1$ then $a_t \to 0$.*
2. *If $\beta > 1$ then $|a_t| \to \infty$.*

Proof. First, observe that (A1) and (A5) imply

$$\begin{aligned} a_t - \frac{\gamma}{1-\beta} &= a_{t-1} - \frac{\gamma}{1-\beta} + \phi\left(\mathcal{T}_\gamma\left(a_{t-1},\omega_{t-1},\beta\right) - \frac{\gamma}{1-\beta} - \left(a_{t-1} - \frac{\gamma}{1-\beta}\right)\right) \\ &= a_{t-1} - \frac{\gamma}{1-\beta} + \phi\left(\mathcal{T}_0\left(a_{t-1} - \frac{\gamma}{1-\beta},\omega_{t-1},\beta\right) - \left(a_{t-1} - \frac{\gamma}{1-\beta}\right)\right), \end{aligned}$$

which shows that it suffices to prove the results for $\gamma = 0$. We drop the subscript on T.

Now assume $|\beta| < 1$, and observe that *for any* ω,

$$\left|\beta\sum_{k\geq 0}\beta^k\omega_k\right| \leq |\beta|\sum_{k\geq 0}|\beta^k|\omega_k \leq |\beta|\sum_{k\geq 0}|\beta|\,\omega_k \leq \beta^2. \tag{A6}$$

Next, write the recursion (A5) as

$$a_t = \left(1-\phi\left(1-\beta\sum_{k\geq 0}\beta^k\omega_{kt-1}\right)\right)a_{t-1} \equiv A_{t-1}a_{t-1}.$$

By equation (A6),

$$-1 < 1-\phi(1+\beta^2) \leq A_{t-1} \leq 1-\phi(1-\beta^2) < 1.$$

It follows that

$$|a_t| = \left(\prod_{n=1}^{t}A_{t-n}\right)|a_0| \to 0,$$

establishing item 1.

Now let $\beta > 1$. The same reasoning as in (A6), but with the inequalities reversed, yields

$$\beta \sum_{k\geq 0} \beta^k \omega_k \geq \beta^2.$$

It follows that

$$A_t \geq 1 - \phi + \phi\beta^2 = 1 + \phi(\beta^2 - 1) > 1,$$

and the result follows. ■

Proof of Theorem 1. The result is immediate: since Lemma A.2 holds for any sequence of weight systems, it holds in particular for whatever system of weights is produced by the unified dynamics. ■

Stability of the replicator dynamic. We begin with three lemmas.

Lemma A.3. *Suppose* $\gamma = 0$.

1. *If* $|\beta| < 1$ *then* $k < \hat{k}(y)$ *implies that there exists* $\delta \in (0,1)$ *such that* $|y| < (1-\delta)|a\beta^k|$.
2. *If* $\beta > 1$ *then* $k < \hat{k}(y)$ *implies that there exists* $\delta > 0$ *such* $|y| > (1+\delta)|a\beta^k|$.

Proof. Assume $|\beta| < 1$. If $|y| < |a\beta^{\hat{k}}|$ we are done, so assume $|a\beta^{\hat{k}}| \leq |y|$. Let $\delta = 1/2(1 - |\beta^{\hat{k}-k}|)$. We claim $2|y| < |a\beta^{\hat{k}}| + |a\beta^{\hat{k}-1}|$. Indeed, by the optimality of $\hat{k}$,

$$|y| - |a\beta^{\hat{k}}| = |y - a\beta^{\hat{k}}| < |y - a\beta^{\hat{k}-1}| = |a\beta^{\hat{k}-1}| - |y|.$$

Thus we compute

$$\begin{aligned} |y| &< \frac{1}{2}\left(|a\beta^{\hat{k}}| + |a\beta^{\hat{k}-1}|\right) \leq \frac{1}{2}\left(|a\beta^{\hat{k}}| + |a\beta^{k}|\right) \\ &= \frac{1}{2}\left(|\beta^{\hat{k}-k}| + 1\right)|a\beta^k| = (1-\delta)|a\beta^k|. \end{aligned}$$

Now assume $\beta > 1$. We may also assume, without loss of generality, that $a > 0$. Let $\delta = 1/2(|\beta^{\hat{k}-k}| - 1)$. If $y > a\beta^{\hat{k}}$ we are done, so assume $a\beta^{\hat{k}} \geq y$. It follows that

$$a\beta^{\hat{k}} \geq y > \frac{a}{2}\left(\beta^{\hat{k}} + \beta^k\right) = \frac{1}{2}\left(\beta^{\hat{k}-k} + 1\right)a\beta^k = (1+\delta)a\beta^k,$$

where the second inequality follows from the definition of $\hat{k}$. ■

Lemma A.4. *Let* $\gamma = 0$ *and* $\{y_t\}_{t\geq 1}$ *be generated by the replicator, initialized with weights* $\{\omega_{n0}\}_{n\in\mathbb{N}}$ *and beliefs* a. *Let* $\check{k} \geq 1$ *and suppose there exists* $N > 0$ *such that* $t \geq N$ *implies* $\hat{k}(y_t) > \check{k}$. *Then* $\lim_{t\to\infty} \omega_{nt} = 0$ *for all* $n \leq \check{k}$.

Proof. Let $t \geq N$. First suppose $|\beta| < 1$. Since $\hat{k}(y_t) > \check{k}$, it follows from Lemma A.3 that $(1-\delta)|a\beta^{\check{k}}| > |y_t|$, for some $\delta \in (0,1)$. Thus $n \leq \check{k}$ implies

$$|a\beta^n - y_t| \geq |a\beta^n| - |y_t| > |a\beta^n| - (1-\delta)|a\beta^{\check{k}}| > 0.$$

Using this estimate in the replicator yields, and that $r' > 0$, we have, for $s \geq 1$,

$$\begin{aligned}\omega_{nt+s} &= (1 - r(|a\beta^n - y_{t+s-1}|))\,\omega_{nt+s-1} \\ &< \left(1 - r\left(|a\beta^n| - (1-\delta)|a\beta^{\check{k}}|\right)\right)\omega_{nt+s-1} \\ &< \left(1 - r\left(|a\beta^n| - (1-\delta)|a\beta^{\check{k}}|\right)\right)^s \omega_{nt-1}.\end{aligned}$$

Because $r(0) \geq 0$ it follows that $\omega_{nt+s} \to 0$ as $s \to \infty$.

Now suppose $\beta > 1$, and assume, without loss of generality, that $a > 0$. Since $\hat{k}(y_t) > \check{k}$, it follows from Lemma A.3 that $a\beta^{\check{k}}(1+\delta) < y_t$. Thus $n \leq \check{k}$ implies

$$|a\beta^n - y_t| = y_t - a\beta^n \geq (1+\delta)a\beta^{\check{k}} - a\beta^n > 0.$$

The argument now proceeds analogously to the case $|\beta| < 1$. ■

Lemma A.5. *If x_n is an integer sequence and $\liminf x_n = x < \infty$ then there exists $N > 0$ such that $n \geq N$ implies $x_n \geq x$.*

Proof. The result is trivial if $x = -\infty$ so assume otherwise. Let $\hat{x}_k = \inf_{n \geq k} x_n$. Then $\hat{x}_k$ is a non-decreasing integer sequence converging to x. Now simply choose N so that $|\hat{x}_N - x| < 1$. ■

We are now ready to prove the main result.

Proof of Theorem 2. By Lemma A.1 we may assume $\gamma = 0$. To thin notation, let $\hat{k}_t = \hat{k}(y_t)$. It is helpful to introduce the relation $\succ$: for $y \in \mathbb{R}$ and $m(y), n(y) \in \mathbb{N}$, write $m(y) \succ n(y)$ when the level-m forecast is superior to the level-n forecast, i.e.,

$$m(y) \succ n(y) \iff |y - a\beta^{m(y)}| < |y - a\beta^{n(y)}|.$$

Now set $\tilde{k} = \liminf \hat{k}_t$.

We consider the cases $\beta > 1$ and $|\beta| < 1$ separately, however, we note that for each case it suffices to show $\tilde{k} = \infty$. To see this, first consider the case $\beta > 1$, and note that without loss of generality we may assume $a > 0$. Let $\Delta > 0$ and pick m so that $a\beta^m > \Delta$. Since $\tilde{k} = \infty$ it follows that $\hat{k}_t \to \infty$, so pick $\hat{t}$ so that $t \geq \hat{t} \implies \hat{k}_t > m$. Finally, for $n \geq 1$ let $\Omega_t^l(n) = \sum_{k<n} \omega_{kt}$, and note that, by Lemma A.4, $\tilde{k} = \infty$ implies $\Omega_t^l(n) \to 0$ as $t \to \infty$. Thus

$$\lim_{t\to\infty} y_t = \lim_{t\to\infty} a\beta \sum_{n\in\mathbb{N}} \beta^n \omega_{nt} \geq \lim_{t\to\infty} \left(1 - \Omega_t^l(m)\right) a\beta^{m+1} = a\beta^{m+1} > \Delta.$$

Now suppose $|\beta| < 1$. By Lemma A.4, if $\hat{k}_t \to \infty$ then all the weights are driven to zero. If all the weights are driven to zero then $y_t \to 0$: indeed, writing, $\omega_t^{\max} = \max_{i\in\mathbb{N}} \omega_{it}$, we have

$$|y_t| = \left| a\beta \sum_{n\in\mathbb{N}} \omega_{nt}\beta^n \right| \leq \omega_t^{\max} |a\beta| \sum_{n\in\mathbb{N}} |\beta^n| \to 0,$$

since $\omega_t^{\max} \to 0$ as $t \to \infty$.

Our proof strategy is to assume $\tilde{k} < \infty$ and derive a contradiction. To this end, it suffices to find some $M > 0$ so that $t \geq M$ implies the existence of

$m(y_t) > \tilde{k}$ with $m(y_t) \succ \tilde{k}$, as this contradicts the definition of $\tilde{k}$ as the limit infimum of the $\hat{k}_t$.

First, the easy case: $\beta > 1$; and again assume $a > 0$. Then $y_t \geq (1 - \Omega_t^l(\tilde{k}))a\beta^{\tilde{k}+1}$, and, by Lemmas A.4 and A.5, $\lim_{t\to\infty} \Omega_t^l(\tilde{k})) = 0$. It follows that eventually, $\tilde{k}+1 \succ \tilde{k}$, which is the desired contradiction.

Now, assume $|\beta| < 1$, and let N be chosen as in Lemma A.5. The desired contradiction is developed in three steps.

Step 1. We establish the following claim:

Claim. *Given $\varepsilon > 0$ there exists $\mathcal{M}(\varepsilon) > 0$ so that $t \geq \mathcal{M}(\varepsilon) \geq N$ implies $|y_t| < |a\beta|^{\tilde{k}+1}(1+\varepsilon)$.*

Proof of claim. We know that for all $t \geq N$ we have $\hat{k}_t \geq \tilde{k}$. It follows that, for $\tilde{k} \geq 1$,

$$\begin{aligned} |y_t| &\leq |a\beta| \sum_{k<\tilde{k}} |\beta|^k \omega_{kt} + |a\beta| \sum_{k\geq\tilde{k}} |\beta|^k \omega_{kt} \\ &< |a\beta|\Omega_t^l\left(\tilde{k}\right) + |a||\beta|^{\tilde{k}+1}\left(1 - \Omega_t^l\left(\tilde{k}\right)\right). \end{aligned}$$

By Lemma A.4 we have that $\Omega_t^l\left(\tilde{k}\right) \to 0$ as $t \to \infty$, which establishes the claim.

Step 2. We now prove the result when $0 < \beta < 1$. Choose $2\varepsilon < \beta^{-1} - 1$ so that

$$(1+\varepsilon)\beta^{n+1} < \frac{1}{2}\left(\beta^{n+1} + \beta^n\right).$$

Let $\mathcal{M}(\varepsilon) = \mathcal{M}$ be chosen as in Step 1, and assume $t \geq \mathcal{M}$. There are two cases.

Case 1: $a > 0$. It follows that $y_t > 0$. Then

$$0 < y_t < a\beta^{\tilde{k}+1}(1+\varepsilon) < \frac{1}{2}\left(a\beta^{\tilde{k}+1} + a\beta^{\tilde{k}}\right),$$

which implies that $\tilde{k}+1 \succ \tilde{k}$, the desired contradiction.

Case 2: $a < 0$. In this case $y_t < 0$. Then

$$0 > y_t > a\beta^{\tilde{k}+1}(1+\varepsilon) > \frac{1}{2}\left(a\beta^{\tilde{k}+1} + a\beta^{\tilde{k}}\right),$$

which implies that $\tilde{k}+1 \succ \tilde{k}$, the desired contradiction.

Step 3. Finally, we prove the result when $-1 < \beta < 0$. Choose $\varepsilon < (2|\beta|)^{-1}(1 - |\beta|)^2$ and choose $\mathcal{M}(\varepsilon)$ as in Step 1. Now notice that

$$1 + \varepsilon < (2|\beta|^{(n+1)})^{-1}\left(|\beta|^n + |\beta|^{n+2}\right),$$

for any $n \geq 1$. It follows that

$$\begin{aligned} 2|\beta|^{\tilde{k}+1}(1+\varepsilon) &< |\beta|^{\tilde{k}+2} + |\beta|^{\tilde{k}}, \text{ or} \\ 0 < |\beta|^{\tilde{k}+1}(1+\varepsilon) - |\beta|^{\tilde{k}+2} &< |\beta|^{\tilde{k}} - |\beta|^{\tilde{k}+1}(1+\varepsilon). \end{aligned} \tag{A7}$$

Let $t \geq M$. There are two cases.

Case 1: $\hat{k}_t \neq \tilde{k}$ (mod 2). In this case $\text{sign}(y_t) = -\text{sign}(a\beta^{\tilde{k}})$, whence $\tilde{k}+1 \succ \tilde{k}$.

Case 2: $\hat{k}_t = \tilde{k} \pmod 2$. If $y_t < 0$ then $a\beta^{\tilde{k}}$ is negative. Next, note that if $y_t \geq a\beta^{\tilde{k}+2}$ then $\tilde{k}+2 \succ \tilde{k}$, which is a contradiction. Thus

$$a\beta^{\tilde{k}} < -|a\beta^{\tilde{k}+1}|(1+\varepsilon) < y_t < a\beta^{\tilde{k}+2} < 0,$$

where the first inequality follows from (A7). Thus

$$\begin{aligned} |a\beta^{\tilde{k}+2} - y_t| &< \left| a\beta^{\tilde{k}+2} + |a\beta^{\tilde{k}+1}|(1+\varepsilon) \right| \\ &= |a\beta^{\tilde{k}+1}|(1+\varepsilon) - |a\beta^{\tilde{k}+2}| \\ &= |a| \left(|\beta^{\tilde{k}+1}|(1+\varepsilon) - |\beta^{\tilde{k}+2}| \right) \\ &< |a| \left(|\beta^{\tilde{k}}| - |\beta^{\tilde{k}+1}(1+\varepsilon)| \right) \\ &= |a\beta^{\tilde{k}}| - |a\beta^{\tilde{k}+1}(1+\varepsilon)| \\ &< |a\beta^{\tilde{k}}| - |y_t| = \left| a\beta^{\tilde{k}} - y_t \right|, \end{aligned}$$

which implies $\tilde{k}+2 \succ \tilde{k}$.

Now suppose $y_t > 0$, so that $a\beta^{\tilde{k}}$ is positive. Thus

$$a\beta^{\tilde{k}} > |a\beta^{\tilde{k}+1}|(1+\varepsilon) > y_t > a\beta^{\tilde{k}+2} > 0,$$

where the reasoning is as above. Thus

$$\begin{aligned} |y_t - a\beta^{\tilde{k}+2}| &< \left| |a\beta^{\tilde{k}+1}|(1+\varepsilon) - a\beta^{\tilde{k}+2} \right| \\ &= |a\beta^{\tilde{k}+1}|(1+\varepsilon) - |a\beta^{\tilde{k}+2}| \\ &= |a| \left(|\beta^{\tilde{k}+1}|(1+\varepsilon) - |\beta^{\tilde{k}+2}| \right) \\ &< |a| \left(|\beta^{\tilde{k}}| - |\beta^{\tilde{k}+1}(1+\varepsilon)| \right) \\ &= |a\beta^{\tilde{k}}| - |a\beta^{\tilde{k}+1}(1+\varepsilon)| \\ &< |a\beta^{\tilde{k}}| - |y_t| = \left| a\beta^{\tilde{k}} - y_t \right|, \end{aligned}$$

so that $\tilde{k}+2 \succ \tilde{k}$, completing the proof of step 3. ∎

Proof of Theorem 3. Lemma A.2 establishes items 1 and 2, and so we focus here only on item 3. Also, as demonstrated in the proof of Lemma A.2, we may assume $\gamma = 0$. We recall the notation $\Omega = \dot{\cup}_n \Delta^n$ and $\psi_\beta : \Omega \to \mathbb{R}$, given by $\psi_\beta(\omega) = \beta \sum_k \beta^k \omega_k$, and that Ω is endowed with the direct-limit topology.

The dynamic system for a_t may be written

$$a_t = (1 - \phi + \phi\psi_\beta(\omega))\, a_{t-1} \equiv A(\beta, \omega, \phi) a_{t-1}.$$

It follows that $|A(\beta,\omega,\phi)| < 1 \implies a_t \to 0$ and $|A(\beta,\omega,\phi)| > 1 \implies |a_t| \to \infty$. We compute

$$\begin{aligned} &|A(\beta,\omega,\phi)| < 1 \iff -1 < 1 - \phi + \phi\psi_\beta(\omega) < 1 \iff 1 - 2\phi^{-1} < \psi_\beta(\omega) < 1, \text{ and} \\ &|A(\beta,\omega,\phi)| > 1 \iff 1 - \phi + \phi\psi_\beta(\omega) < -1 \text{ or } 1 - \phi + \phi\psi_\beta(\omega) > 1 \\ &\qquad \iff \psi_\beta(\omega) < 1 - 2\phi^{-1} \text{ or } \psi_\beta(\omega) > 1. \end{aligned}$$

This completes the proof of items 3(a) - 3(c).

To establish item 3(d) we start by showing that ψ_β is continuous. Let ψ_β^n be the restriction of ψ_β to $\Delta^n \subset \Omega$. It suffices to show that $\psi_\beta^n : \Delta^n \to \mathbb{R}$ is continuous for each $n \in \mathbb{N}$. To see this, let $U \subset \mathbb{R}$ be open. Then

$$\psi_\beta^{-1}(U) = \cup_n \left(\psi_\beta^{-1}(U) \cap \Delta^n\right) = \cup_n \left(\left(\psi_\beta^n\right)^{-1}(U) \cap \Delta^n\right) = \cup_n \left(\left(\psi_\beta^n\right)^{-1}(U)\right).$$

Assuming $\psi_\beta^n : \Delta^n \to \mathbb{R}$ is continuous, we have that $\left(\psi_\beta^n\right)^{-1}(U)$ is open in Δ^n, whence open in Ω. Thus $\psi_\beta^{-1}(U)$ is a union of open sets in Ω, which establishes the continuity of ψ_β.

Next we demonstrate surjectivity of ψ_β. Let $z \in \mathbb{R}$. Since $\beta < -1$ we can find an $n \in \mathbb{N}$ with $n \geq 1$ so that $\beta^{2n+1} < z < \beta^{2n}$. By continuity there is $\varepsilon \in (0, 1/2)$ such that

$$(1-\varepsilon)\beta^{2n+1} + \varepsilon\beta^{2n} < z < \varepsilon\beta^{2n+1} + (1-\varepsilon)\beta^{2n}.$$

For $\alpha \in (0,1)$ let $\omega(\alpha) \in \Delta^{2n+1} \subset \Omega$ be given by

$$\omega_k(\alpha) = \begin{cases} \alpha & \text{if } k = 2n+1 \\ 1-\alpha & \text{if } k = 2n \\ 0 & \text{else} \end{cases}$$

and note that $\alpha \to \omega_\alpha$ continuously maps $(0,1)$ into Δ^{2n+1}, whence into Ω. Let $\Psi_\beta : (0,1) \to \mathbb{R}$ be $\Psi_\beta(\alpha) = \psi_\beta(\omega(\alpha))$. It follows that Ψ_β is continuous and

$$\Psi_\beta(\varepsilon) = (1-\varepsilon)\beta^{2n+1} + \varepsilon\beta^{2n} < z < \varepsilon\beta^{2n+1} + (1-\varepsilon)\beta^{2n} = \Psi_\beta(1-\varepsilon)$$

By the intermediate value theorem there is an $\alpha \in (\varepsilon, 1-\varepsilon)$ so that $z = \Psi_\beta(\alpha) = \psi_\beta(\omega(\alpha))$, which establishes surjectivity.

Now let

$$\begin{aligned} \Omega_s &= \psi_\beta^{-1}\left((1-2\phi^{-1}, 1)\right) \\ \Omega_u &= \psi_\beta^{-1}\left((-\infty, 1-2\phi^{-1}) \cup (1,\infty)\right). \end{aligned}$$

Both sets are open by the continuity of ψ_β, and from items 3(a) and 3(b) we have that $\omega \in \Omega_s$ implies $y_t \to \bar{y}$ and $\omega \in \Omega_u$ implies $|y_t| \to \infty$. Thus parts (i) and (ii) of item 3(d) are established.

Finally, let $\Omega_0 = \Omega \setminus (\Omega_s \cup \Omega_u)$. We must show that Ω_0 is no-where dense, i.e. that the interior of the closure of Ω_0 is empty. To this end, notice that

$$\Omega_0 = \psi_\beta^{-1}(\{-1\}) \,\dot\cup\, \psi_\beta^{-1}(\{1\}) \equiv \Omega_0^- \,\dot\cup\, \Omega_0^+.$$

Since ψ_β is continuous, it follows that $\Omega_0^\pm$ are closed. Since no-where denseness is closed under finite unions, it suffices to show that the interiors of $\Omega_0^\pm$ are empty. Thus let $\omega \in \Omega_0^+$. Let $N \in \mathbb{N}$ so that $\omega \in \Delta^N$. Since $\beta < -1$ and $\psi_\beta(\omega) = 1$ there is an even $n \in \mathbb{N}$ and an odd $m \in \mathbb{N}$, with $n, m \leq N$ and such that $\omega_n, \omega_m \neq 0$. For $k \in \mathbb{N}$ with $k \geq 2$, define $\omega^k \in \Delta^N \subset \Omega$ as follows:

$$\omega_i^k = \begin{cases} (1-k^{-1})\omega_n & \text{if } i = n \\ \omega_m + k^{-1}\omega_n & \text{if } i = m \\ \omega_i & \text{else} \end{cases}$$

Note that ω^k is the same weight system as ω except that some of the weight associated with the positive forecast β^n is shifted to the negative forecast β^m.

Because the model itself has negative feedback, this means that the implied value of y is larger for weight system ω^k than it is for weight system ω. More formally, $k \geq 2$ implies that $\psi_\beta(\omega^k) > 1$, which implies that $\omega^k \in \Omega_u$. Now notice that, as a sequence in Δ^N, we have $\omega^k \to \omega$. Owing to the construction of the direct-limit topology, we have that $\omega^k \to \omega$ in Ω as well. Thus, given an arbitrary element $\omega \in \Omega_0^+$ we have constructed a sequence in Ω_u converging to it, and since $\Omega_u \cap \Omega_0^+$ is empty, we conclude that ω is not in the interior of Ω_0^+. So the interior of Ω_0^+ is empty, and since Ω_0^+ is closed, we conclude that Ω_0^+ is nowhere dense. The same argument applies to Ω_0^-, which shows that $\Omega_0 = \Omega_0^- \dot{\cup} \Omega_0^+$ is no-where dense. ■

Full statement and proof of Proposition 1. Recall from (4) that $\hat{k}$ is defined explicitly as a function of y_t. However, both y_t and $E_{t-1}^k y_t$ are affine functions of level-0 beliefs a. In particular, if $\gamma = 0$ then

$$\hat{k}(a) = \min \arg \min_{k \in \mathbb{N}} |\beta^k a - \beta \textstyle\sum_k \omega_k a|, \tag{A8}$$

which further implies that $\hat{k}$ is independent of a. It is straightforward to show this result continues to hold with $\gamma \neq 0$, and, in fact, $\hat{k}$ is independent of the value of γ. Thus, we may view $\hat{k} = \hat{k}(\beta, \omega)$. We have the following result.

Proposition 1′ (Optimal forecast levels)**.** *Let $K \geq 1$ and $\omega^K = \{\omega_n\}_{n=0}^K$ be a weight system with weights given as $\omega_n = (K+1)^{-1}$. Let $\hat{k} = \hat{k}\left(\beta, \omega^K\right)$.*

1. *Suppose $0 < \beta < 1$.*
   (a) $K \to \infty \implies \hat{k} \to \infty$ *and* $\hat{k}/K \to 0$.
   (b) $\beta \to 1^- \implies \hat{k} \to \begin{cases} \frac{K}{2} + 1 & \text{if } K \text{ is even} \\ \frac{K+1}{2} & \text{if } K \text{ is odd} \end{cases}$
   (c) $\beta \to 0^+ \implies \hat{k} \to \begin{cases} 1 & \text{if } K = 1 \\ 2 & \text{if } K \geq 2 \end{cases}$
2. *Suppose $-1 < \beta < 0$.*
   (a) $K \to \infty \implies \hat{k} \to \infty$ *and* $\hat{k}/K \to 0$.
   (b) $\beta \to 0^- \implies \hat{k} \to \begin{cases} 1 & \text{if } K = 1 \\ 3 & \text{if } K \geq 2 \end{cases}$
   (c) $\beta \to -1^+ \implies \hat{k} \to \infty$.
3. *Suppose $\beta < -1$*
   (a) $K \to \infty \implies \hat{k} \to \infty$ *and* $\hat{k}/K \to 1$
   (b) $\beta \to -1^- \implies \hat{k} \to \begin{cases} 1 & \text{if } K \text{ is even} \\ 0 & \text{if } K \text{ is odd} \end{cases}$
   (c) $\beta \to -\infty \implies \hat{k} \to K + 1$.

Before proceeding to the proof, some preliminary work is required. By Lemma A.1 we may assume $\gamma = 0$ and $a = 1$. Recall from Section 3.2 our notation for uniform weights: for $K \in \mathbb{N}$, $\omega^K = \{\omega_n\}_{n=0}^K$ with $\omega_n = (K+1)^{-1}$. It follows that

$$y = \beta \textstyle\sum_k \beta^k \omega_k = \frac{\beta}{K+1} \textstyle\sum_k \beta^k = \frac{\beta\left(1-\beta^{K+1}\right)}{(K+1)(1-\beta)} \equiv \psi\left(K, \beta\right).$$

When it does not impede clarity, we make the identifications $\hat{k} = \hat{k}\left(\beta, \omega^K\right)$ and $\psi = \psi\left(K, \beta\right)$.

It is helpful to define k^* as the *continuous counterpart* to $\hat{k}$. For $\beta > 0$ our definition for k^* corresponds to the first order condition for minimizing $(\beta^k - \psi(K,\beta))^2$ for $k \in \mathbb{R}_+$. However, care must be taken to accommodate $\beta < 0$. We define k^* as follows:

$$k^*\left(K, \beta\right) = \frac{\log\left(\psi\left(K,\beta\right)^2\right)}{\log(\beta^2)}. \tag{A9}$$

Of course if β, and hence ψ, are positive then we can dispense with the squared terms in the definition.

Now define $\lfloor\cdot\rfloor$ to be the usual floor function, i.e. for $x \in \mathbb{R}$, $\lfloor x \rfloor$ is the largest integer less than or equal to x. Define $\lfloor\cdot\rfloor_{\text{odd}}$ and $\lfloor\cdot\rfloor_{\text{even}}$ and the odd and even floors, respective, which take the obvious meaning, e.g. $\lfloor x\rfloor_{\text{even}}$ is the largest even integer less than or equal to x. Finally, $\lceil\cdot\rceil$, $\lceil\cdot\rceil_{\text{even}}$, and $\lceil\cdot\rceil_{\text{even}}$ have the analogous definitions. Define

$$k^*_{\text{low}} = \begin{cases} \lfloor k^* \rfloor & \text{if } 0 < \beta < 1 \\ \lfloor k^* \rfloor_{\text{odd}} & \text{if } -1 < \beta < 0 \text{ or if } \beta < -1 \textit{ and } \psi < \frac{1+\beta}{2} \\ \lfloor k^* \rfloor_{\text{even}} & \text{if } \beta < -1 \textit{ and } \psi > 0 \end{cases}$$

and define k^*_{high} analogously using the ceiling functions. The following result links k^* and $\hat{k}$.

Lemma A.6. *If $k^* \geq 0$ then $\hat{k} \in \{k^*_{low}, k^*_{high}\}$.*

Proof. We begin with the following observations on the parity of $\hat{k}$.[1] Recall that 0 is taken as even.

1. If $-1 < \beta < 0$ then $\hat{k}$ is odd.
2. If $\beta < -1$ and $\psi < \frac{1+\beta}{2}$ then $\hat{k}$ is odd.
3. If $\beta < -1$ and $\psi > 0$ then $\hat{k}$ is even.

These items may be established as follows. Note that $-1 < \beta < 0$ implies $\psi < 0$, whence there is an odd $n \in \mathbb{N}$ so that $\psi < \beta^n < 0$, making n superior to any even forecast level. If $\beta < -1$ and $\psi < \frac{1+\beta}{2}$ then the level 1 forecast is superior to any even forecast level. If $\beta < -1$ and $\psi > 0$ then the level 0 forecast is superior to any odd forecast level.

Next, note that $k^* < 0$ if and only if $-1 < \psi < 1$ and $\beta < -1$. Now, for $\alpha \in \mathbb{R}_+$ define $\phi(\alpha, \beta)$ as follows:

$$\phi(\alpha,\beta) = \begin{cases} \left(\beta^2\right)^{\frac{\alpha}{2}} & \text{if } \psi > 0 \\ \beta\left(\beta^2\right)^{\frac{\alpha-1}{2}} & \text{if } \psi < 0 \end{cases}$$

This function has the following properties:

(a) If $\hat{k} \geq 1$ and if non-zero $k \in \mathbb{N}$ has the same parity as $\hat{k}$ then $\beta^k = \phi(k, \beta)$: in this way ϕ extends our notion of forecast level to all positive reals.

(b) $\phi(k^*, \beta) = \psi$.

[1] The *parity* of $n \in \mathbb{N}$ is its equivalence class mod 2. Thus n and m have the same parity if they are either both even or both odd.

To establish item (a), first suppose $\hat{k}$ is even. Since $\hat{k} \geq 1$ it follows that $\psi > 0$. Let $k = 2m$ for $m > 0$. Then $\phi(k, \beta) = (\beta^2)^m = \beta^k$. Next suppose $\hat{k}$ is odd. Let $k = 2m + 1$. If $0 < \beta < 1$ then $\psi > 0$, so that $\phi(k, \beta) = (\beta^2)^{\frac{2m+1}{2}} = \beta^{2m+1}$. Let $\beta < 0$. If $-1 < \beta < 0$ then $\psi < 0$. If $\beta < -1$ then $\hat{k}$ odd implies $\psi < 0$. Thus $k = 2m + 1$ implies $\phi(k, \beta) = \beta(\beta^2)^m = \beta^{2m+1}$. To establish item (b), observe that $\psi > 0$ implies

$$\log \phi(k^*, \beta) = ({k^*}/{2}) \log \beta^2 = ({1}/{2}) \log \psi^2 = \log \psi$$

and $\psi < 0$ implies $\phi(k^*, \beta) < 0$, and

$$\log\left(-\phi(k^*, \beta)\right) = \log\left(\beta^2\right)^{\frac{1}{2}}\left(\beta^2\right)^{\frac{k^*-1}{2}} = \log\left(\beta^2\right)^{\frac{k^*}{2}} = ({k^*}/{2}) \log \beta^2 = \log\left(-\psi\right).$$

We turn now to the body of the proof of Lemma A.6, in which we use the following notation: $k_1 \prec k_2$ if β^{k_1} is strictly inferior to β^{k_2} as a forecast of ψ. The strategy is as follows: show that $k < \lfloor k^* \rfloor \implies k \prec \lfloor k^* \rfloor$, and that $k > \lceil k^* \rceil$ implies that $k \prec \lceil k^* \rceil$, with floor and ceiling functions adjusted for parity as needed.

Case 1: $0 < \beta < 1$. Since $\psi < \beta$ in this case, we have that $k^* \geq 1$ and $\hat{k} \geq 1$. Also $\alpha > 0$ implies $\phi_\alpha(\alpha, \beta) < 0$. Thus if $k_1 < \lfloor k^* \rfloor$ and $k_2 > \lceil k^* \rceil$ then

$$\phi(k_1, \beta) > \phi(\lfloor k^* \rfloor, \beta) \geq \underbrace{\phi(k^*, \beta)}_{\psi} \geq \phi(\lceil k^* \rceil, \beta) > \phi(k_2, \beta).$$

Thus $k_1 \prec \lfloor k^* \rfloor$ and $k_2 \prec \lceil k^* \rceil$.

Case 2: $-1 < \beta < 0$. Since $\beta < \psi < 0$ in this case, we have that $k^* \geq 1$. Also $\alpha > 0$ implies $\phi_\alpha(\alpha, \beta) > 0$. Also $\psi < 0$ so that $\hat{k}$ is necessarily odd. Thus if $\lfloor k^* \rfloor_{\text{odd}} \geq 1$ and if k_i are odd with $k_1 < \lfloor k^* \rfloor_{\text{odd}}$ and $k_2 > \lceil k^* \rceil_{\text{odd}}$, then

$$\phi(k_1, \beta) < \phi(\lfloor k^* \rfloor_{\text{odd}}, \beta) \leq \underbrace{\phi(k^*, \beta)}_{\psi} \leq \phi(\lceil k^* \rceil_{\text{odd}}, \beta) < \phi(k_2, \beta).$$

Thus $k_1 \prec \lfloor k^* \rfloor_{\text{odd}}$ and $k_2 \prec \lceil k^* \rceil_{\text{odd}}$.

Case 3: $\beta < -1$ and $\psi < \frac{1+\beta}{2}$. Then $k^* \geq 1$ and $\hat{k}$ is odd. Also $\alpha > 1$ implies $\phi_\alpha(\alpha, \beta) < 0$. Thus if $\lfloor k^* \rfloor_{\text{odd}} > 1$ and if k_i are odd with $k_1 < \lfloor k^* \rfloor_{\text{odd}}$ and $k_2 > \lceil k^* \rceil_{\text{odd}}$, then

$$\phi(k_1, \beta) > \phi(\lfloor k^* \rfloor_{\text{odd}}, \beta) \geq \underbrace{\phi(k^*, \beta)}_{\psi} \geq \phi(\lceil k^* \rceil_{\text{odd}}, \beta) > \phi(k_2, \beta).$$

Thus $k_1 \prec \lfloor k^* \rfloor_{\text{odd}}$ and $k_2 \prec \lceil k^* \rceil_{\text{odd}}$.

Case 4: $\beta < -1$ and $\psi > 0$. Then $k^* \geq 0$ (by assumption) and $\hat{k}$ is even. Also $\alpha > 0$ implies $\phi_\alpha(\alpha, \beta) > 0$. Thus if $\lfloor k^* \rfloor_{\text{even}} > 2$ and if k_i are even with $k_1 < \lfloor k^* \rfloor_{\text{even}}$ and $k_2 > \lceil k^* \rceil_{\text{even}}$, then

$$\phi(k_1, \beta) < \phi(\lfloor k^* \rfloor, \beta) \leq \underbrace{\phi(k^*, \beta)}_{\psi} \leq \phi(\lceil k^* \rceil_{\text{even}}, \beta) < \phi(k_2, \beta).$$

Thus $\lfloor k^* \rfloor_{\text{even}} > 2$ implies $k_1 \prec \lfloor k^* \rfloor_{\text{even}}$ and $k_2 \prec \lceil k^* \rceil_{\text{even}}$. If $\lfloor k^* \rfloor_{\text{even}} = 2$ then

$$1 \equiv \beta^0 < \beta^2 = \phi(\lfloor k^* \rfloor_{\text{even}}, \beta) \leq \underbrace{\phi(k^*, \beta)}_{\psi} \leq \phi(\lceil k^* \rceil_{\text{even}}, \beta) < \phi(k_2, \beta).$$

If $\lfloor k^* \rfloor_{\text{even}} = 0 < k^*$ then

$$1 \equiv \beta^0 < \underbrace{\phi(k^*, \beta)}_{\psi} \leq \phi(\lceil k^* \rceil_{\text{even}}, \beta) < \phi(k_2, \beta).$$

Finally, if $k^* = 0$ then $\hat{k} = k^*$. ■

We now turn to the proof of Proposition 1. We note that if $K = 0$ then $k^* = \hat{k} = 1$ regardless of the value of β, so this case is excluded.

Proof of Proposition 1. The arguments for the limits involving $K \to \infty$ will rely directly on the behavior of k^*. The arguments involving limits in β require additional analysis. Define

$$\Delta(k_1, k_2, \beta) = \left(\beta^{k_1} - \psi(\beta)\right)^2 - \left(\beta^{k_2} - \psi(\beta)\right)^2,$$

and note that $k_1 \prec k_2$ when $\Delta(k_1, k_2, \beta) > 0$ and $k_2 \prec k_1$ when $\Delta(k_1, k_2, \beta) < 0$, where the ordering here is as defined in the proof of Lemma A.6. The proof strategy for limiting values of β has three steps:

1. Compute the relevant limiting value of k^*.
2. Use Lemma A.6 to determine a finite set $\hat{\mathcal{K}}$ of possible limiting values for $\hat{k}$.
3. Expand Δ around the limiting value of β and use the expansion to pairwise compare the elements of the $\hat{\mathcal{K}}$.

A final comment before proceeding: Many of the arguments below include tedious symbolic manipulation, and we have relegated much of this work to Mathematica. Whenever Mathematica is relied upon to reach a conclusion, we state this reliance explicitly. As an example, the code used for the first result is included below. All code is available upon request.

Case 1: $0 < \beta < 1$. The following Mathematica code establishes that $K \to \infty$ implies $k^* \to \infty$ and $k^*/K \to 0$.

```
psi[K_, beta_] := beta/(K + 1) Sum[beta^(k - 1), {k, 1, K + 1}];
kstar[K_, beta_] := Log[psi[K, beta]^2]/Log[beta^2];
Module[{limK, limKk, assume},
  assume = {0 < beta < 1};
  limK = Limit[kstar[K, beta], K -> \[Infinity], Assumptions -> And @@ assume];
  limKk = Limit[kstar[K, beta]/K, K -> \[Infinity], Assumptions -> And @@ assume];
  Print["Limit of kstar as K -> infinity is " <> ToString@limK];
  Print["Limit of kstar/K as K -> infinity is " <> ToString@limKk];
  ];
```

Lemma A.6 then implies the same limits for $\hat{k}$, thus proving item 1(a).

Turning to item 1(b), using Mathematica, we find that $\beta \to 1^-$ implies $k^* \to K/2 + 1$. Suppose K is odd. It follows that β near (and below) 1 implies $\lfloor k^* \rfloor < k^* < \lceil k^* \rceil$, whence

$$\hat{k} \in \{\lfloor k^* \rfloor, \lceil k^* \rceil\} = \left\{\frac{K+1}{2}, \frac{K+3}{2}\right\}.$$

Using Mathematica, we find that near $\beta = 1$,

$$\Delta\left(\frac{K+1}{2}, \frac{K+3}{2}, \beta\right) = \frac{1}{12}(K-1)(K+3)(\beta-1)^3 + \mathcal{O}\left(|\beta - 1|^4\right),$$

so that when $K \geq 3$ and β is near and below 1, we conclude that $\Delta < 0$, so that $\hat{k} = 1/2(K+1)$. When $K = 1$ a direct computation shows $\Delta = 0$, so that both $\lfloor k^* \rfloor$ and $\lceil k^* \rceil$ yield the same forecast. Our tiebreaker, then, chooses $\hat{k} = 1$.

Now suppose K is even. Then for β near and below 1 we know that k^* is near $K/2 + 1 \in \mathbb{N}$. Unfortunately, we do not know if k^* approaches its limit monotonically. Thus we can only conclude that for β near and below 1 we have

$$\hat{k} \in \left\{\frac{K}{2}, \frac{K+2}{2}, \frac{K+4}{2}\right\}.$$

Using Mathematica, we find that near $\beta = 1$,

$$\Delta\left(\frac{K}{2}, \frac{K+2}{2}, \beta\right) = (\beta-1)^2 + \mathcal{O}\left(|\beta-1|^3\right)$$
$$\Delta\left(\frac{K+2}{2}, \frac{K+4}{2}, \beta\right) = -(\beta-1)^2 + \mathcal{O}\left(|\beta-1|^3\right).$$

It follows that near and below $\beta = 1$ we have $\frac{K}{2}, \frac{K+4}{2} \prec \frac{K+2}{2}$.
For item 1(c), using Mathematica, we find that $\beta \to 0^+$ implies $k^* \to 1$, so that for small positive β, $\hat{k} \in \{1, 2\}$. Also, $\beta \to 0^+ \implies \psi \to 0$, so $\hat{k} \neq 0$. Using Mathematica, we find that near $\beta = 0$,

$$\Delta(1, 2, \beta) = (2 - 4(1+K)^{-1})(\beta-1)^2 + \mathcal{O}\left(|\beta-1|^3\right), \tag{A10}$$

so that $\hat{k} = 2$ for $K \geq 2$. When $K = 1$ we again find $\Delta = 0$, so that $\hat{k} = 1$.
Case 2: $-1 < \beta < 0$. We establish item 2(a) by direct analysis, and noting that it suffices to study the behavior of k^*. Noting that $-1 < \psi < 0$, we compute

$$\log\psi^2 = 2\log(-\psi) = \log\left(\frac{\beta}{\beta-1}\right) + \log\left(1-\beta^{K+1}\right) - \log(1+K) \to -\infty \tag{A11}$$

$$K^{-1}\log\psi^2 = K^{-1}\log\left(\frac{\beta}{\beta-1}\right) + K^{-1}\log\left(1-\beta^{K+1}\right) - K^{-1}\log(1+K) \to 0 \tag{A12}$$

Since $k^* = \log\psi^2/\log\beta^2$ and $\log\beta^2 < 0$ we see that by equation (A11) $k^* \to \infty$, and that by equation (A12) $k^*/K \to 0$.

Turning to item 2(b), using Mathematica we find that $\beta \to 0^-$ implies $k^* \to 1$, and since $\beta \in (0, 1)$, we know that $\psi < 0$ so that $\hat{k}$ is odd. It follows that for β is near and below 0 we have $\hat{k} \in \{1, 2\}$. The expansion (A10) then shows that $\hat{k} = 3$ for $K \geq 2$. Also as before, $K = 1$ implies $\Delta = 0$, so that $\hat{k} = 1$. Finally, for item 2(c), we find using we find that $\beta \to -1^+$ implies $k^* \to \infty$, and the result follows.
Case 3: $\beta < -1$. We establish item 3(a) by direct analysis. First, observe that $\beta < -1$ implies

$$|\psi(K,\beta)| = \left(\frac{\beta}{\beta-1}\right)\left(\frac{(\beta^2)^{\frac{K+1}{2}} + (-1)^{K+1}}{K+1}\right)$$

By L'Hopital's rule, the function $f(x) = (2\alpha)^{-1}(x^\alpha + \beta)$ diverges to infinity as $\alpha \to \infty$ for $x > 1$ and for any $\beta \in \mathbb{R}$, which shows that $|\psi(K,\beta)| \to \infty$ as $K \to \infty$. It follows that $\log\psi^2 \to \infty$, and thus k^* and $\hat{k}$ go to infinity as $K \to \infty$. Next note

$$\frac{k^*}{K} = (\log(-\beta)^{-1}\left(K^{-1}\log(\beta-1)^{-1}\beta + K^{-1}\log\left((\beta^2)^{\frac{K+1}{2}} + (-1)^{K+1}\right) - K^{-1}\log(K+1)\right).$$

It follows that

$$\lim_{K\to\infty}\frac{k^*}{K} = \lim_{K\to\infty}(K\log(-\beta))^{-1}\log\left((\beta^2)^{\frac{K+1}{2}} + (-1)^{K+1}\right). \tag{A13}$$

Let $g(x) = \alpha^{-1}\log(x^{\frac{\alpha-1}{2}} + \beta)$, for $\beta \in \mathbb{R}$ and $x > 1$. Then

$$\lim_{\alpha\to\infty} g(x) = \lim_{\alpha\to\infty} \frac{x^{\frac{\alpha-1}{2}}\log(x)}{2\left(x^{\frac{\alpha-1}{2}} + \beta\right)} = \log(x)/2.$$

It follows that

$$K^{-1}\log\left((\beta^2)^{\frac{K+1}{2}} + (-1)^{K+1}\right) \to \log(\beta^2)/2 = \log(-\beta),$$

which, when combined with (A13), yields the result.

Turning now to item 3(b), note that if K is odd then $\psi \to 0$, so that $\hat{k} \to 0$. If K is even then $\psi \to -(K+1)^{-1} \in (0,1)$, so that $\hat{k} \to 1$. Finally, for item 3(c), using Mathematica, we find that $\beta \to -\infty$ implies $k^* \to K+1$. By Lemma A.6 we know

$$\lim_{\beta\to-\infty} \hat{k} \in \{K-1, K+1, K+3\}.$$

Again using Mathematica we find that if $K \geq 2$ then

$$\lim_{\beta\to-\infty} \Delta(K-1, K+1) = \lim_{\beta\to-\infty} \Delta(K+1, K+3) - \infty,$$

so that eventually $K+1, K+3 \prec K-1$. If $K = 1$, then $\Delta(K-1, K+1) = 0$ and so by our tie-breaker, $\hat{k} = 0$. ■

A2 Additional Experimental Results

Figure A1 shows the average price observed across all treatments relative to the REE price. Figure A2 shows the individual price predictions for all individuals with outliers indicated by X's. The individual forecasts illustrate both the diversity and uniformity that can occur depending on the expectational feedback in the market. As predicted by the simulations shown in Section 3.4, all the $|\beta| < 1$ cases show convergence to the REE initially and after the announcements, whereas both convergence and non-convergence is observed when $\beta < -1$.

We observed more outliers in individual predictions in this study than were observed, for example, in Bao and Duffy (2016). However, we also have more than double the participants. Some outliers are easily explained as "fat finger" errors where an extra zero is added to a forecast. Others reflect participants with a penchant for anarchy who consistently typed in nonsensical forecasts. In fact, we identify two anarchists who repeatedly typed in the highest price permitted just to see what would happen. One of these anarchists actually provided a nice natural experiment within our laboratory experiment, which we discuss in detail in Section A2.1 below.

When classifying individual forecasts without cutoffs, we chose to not classify 35 out of the 18,367 forecasts from our analysis (5 of which occurred in announcement rounds out of 517 observations in total).[2] Nearly half of the total outliers

[2]The odd number of observations is due to two markets that did not complete the experiment. One market in a T2×A1 treatment ended early when a participant withdrew from the experiment. The other was a T2×A2 treatment that ended a short time after the announcement when a student kicked a power cord knocking out two computers with players in the same market. The data up to that point was saved, but there was no way to let the students pick up where they left off.

Figure A1: Average market price relative to REE

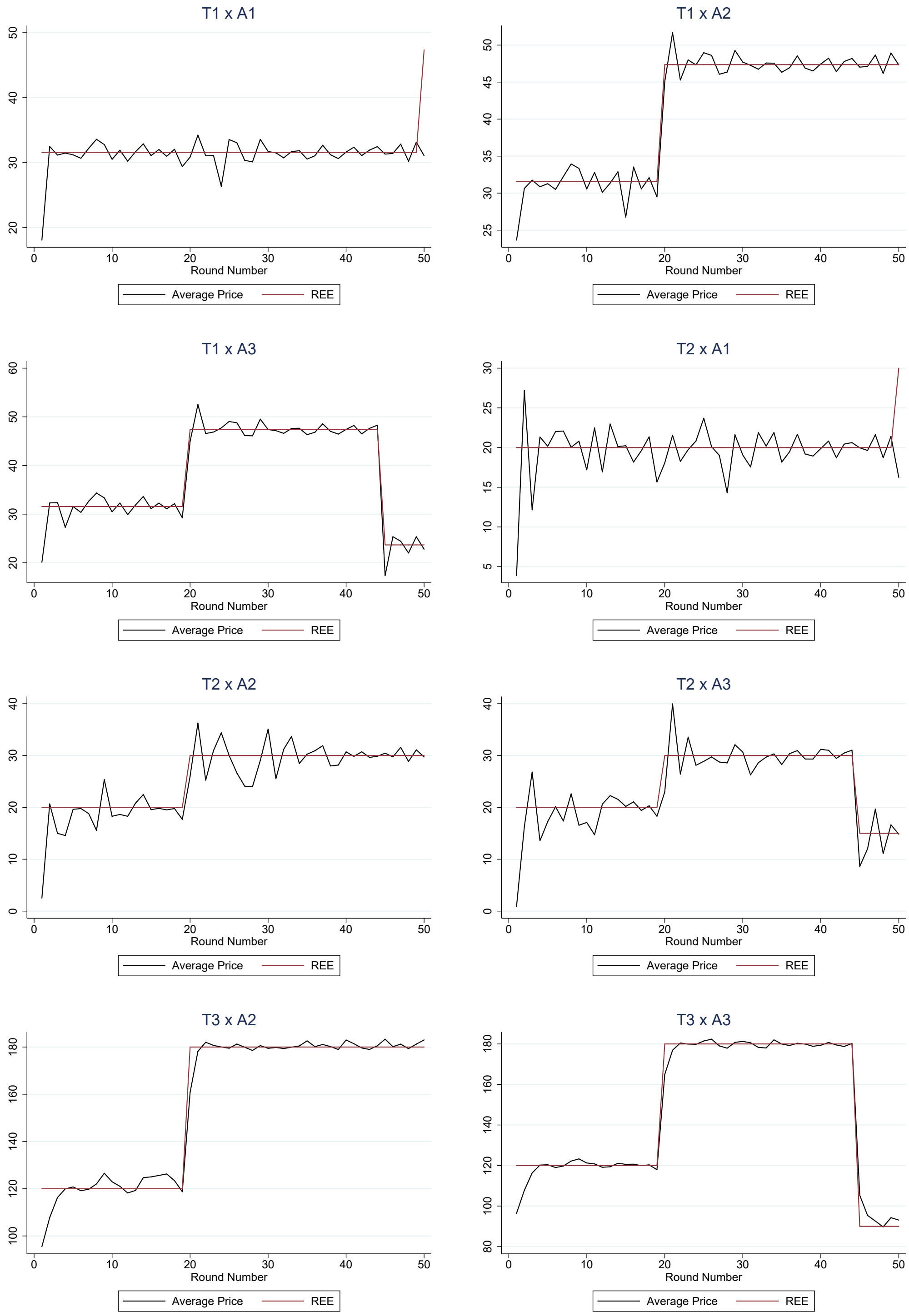
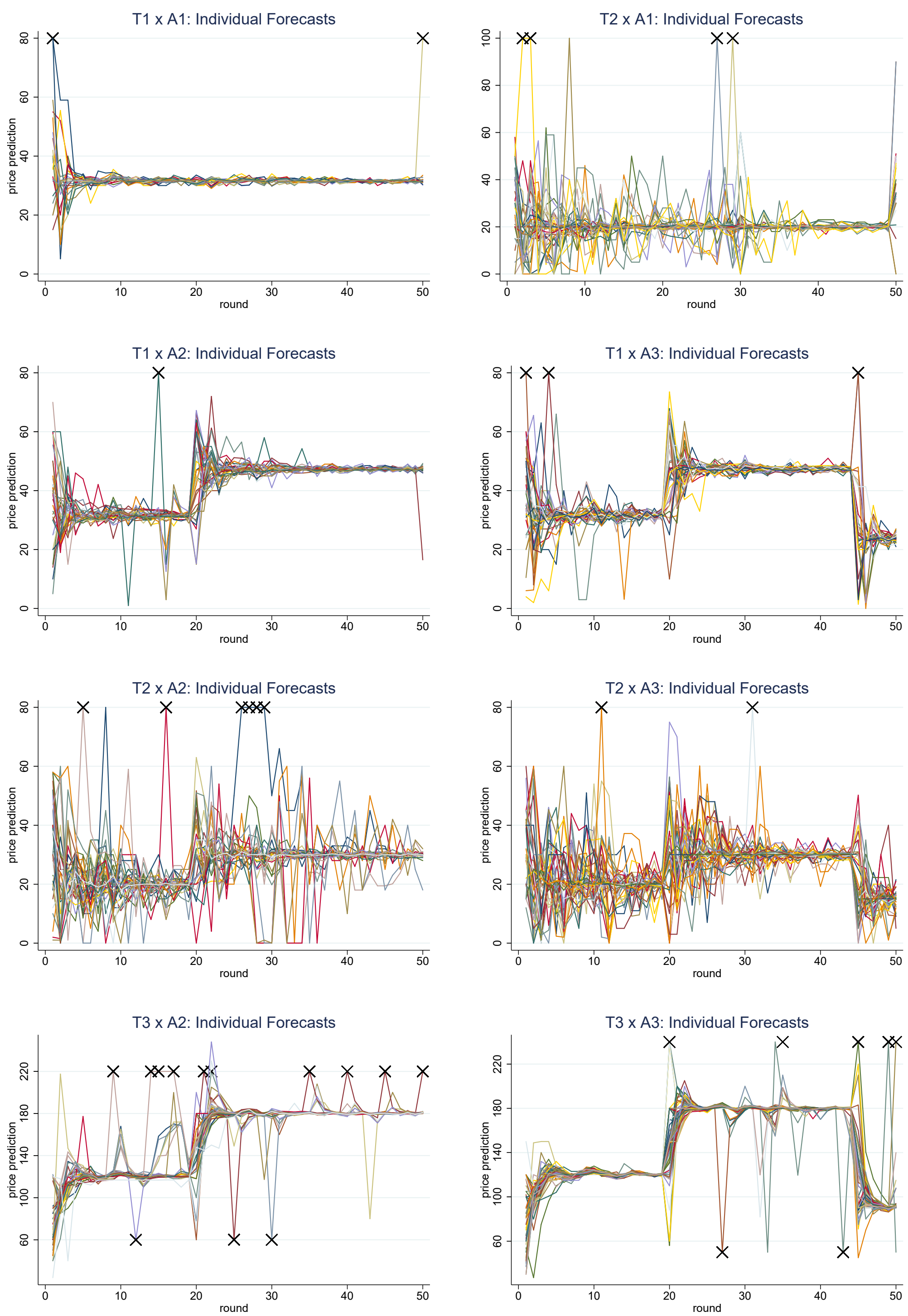


Figure A2: Individual participant predictions



Notes: The 'X's denote forecasts that are larger than the top axis shown in the graph. The maximum value the program would allow a participant to predict is 500.

forecasts were submitted by just 3 (out of the 372) participants in the study. The outlier predictions on average were for a price of 391, which is nearly 200 larger than any plausible price in any treatment. If these outliers were classified as level-k, then most are classified as an REE prediction (e.g. in a positive feedback treatment when level-k forecasts converge from below the REE price and the outlier is above the REE price) or a level-0 prediction (e.g. when convergence starts from above an REE price and level-k deductions are closer to the REE price), which is clearly not in keeping with what the classification is attempting to achieve.[3]

Table A1 provides an overall breakdown of the data, including the outliers, to provide a sense of how far away most forecasts are from the model predictions. The table shows three measures of the root squared difference between a subjects submitted forecast and the nearest level-k model implied forecast, where the level-k forecasts are constructed using the standard assumptions given in Section 5 in the main text. The root mean squared error/difference (RMSE) for the classifications are quite large. This is almost entirely due to outliers and a minority group of the submitted forecasts. The root median squared error/difference (RMedSE) shows that the majority of forecasts are with one unit of a level-k forecast overall and within 4 units in announcement rounds. The final statistic reported in the table is the 70^{th} percentile of root squared differences. This statistic is chosen because we found that approximately 70% of participants chose a level-k forecast in an announcement round when we use a cutoff value of ± 4.5 for pooled data (see Table 4 in the main text). The column illustrates a treatment-by-treatment breakdown of that classification.

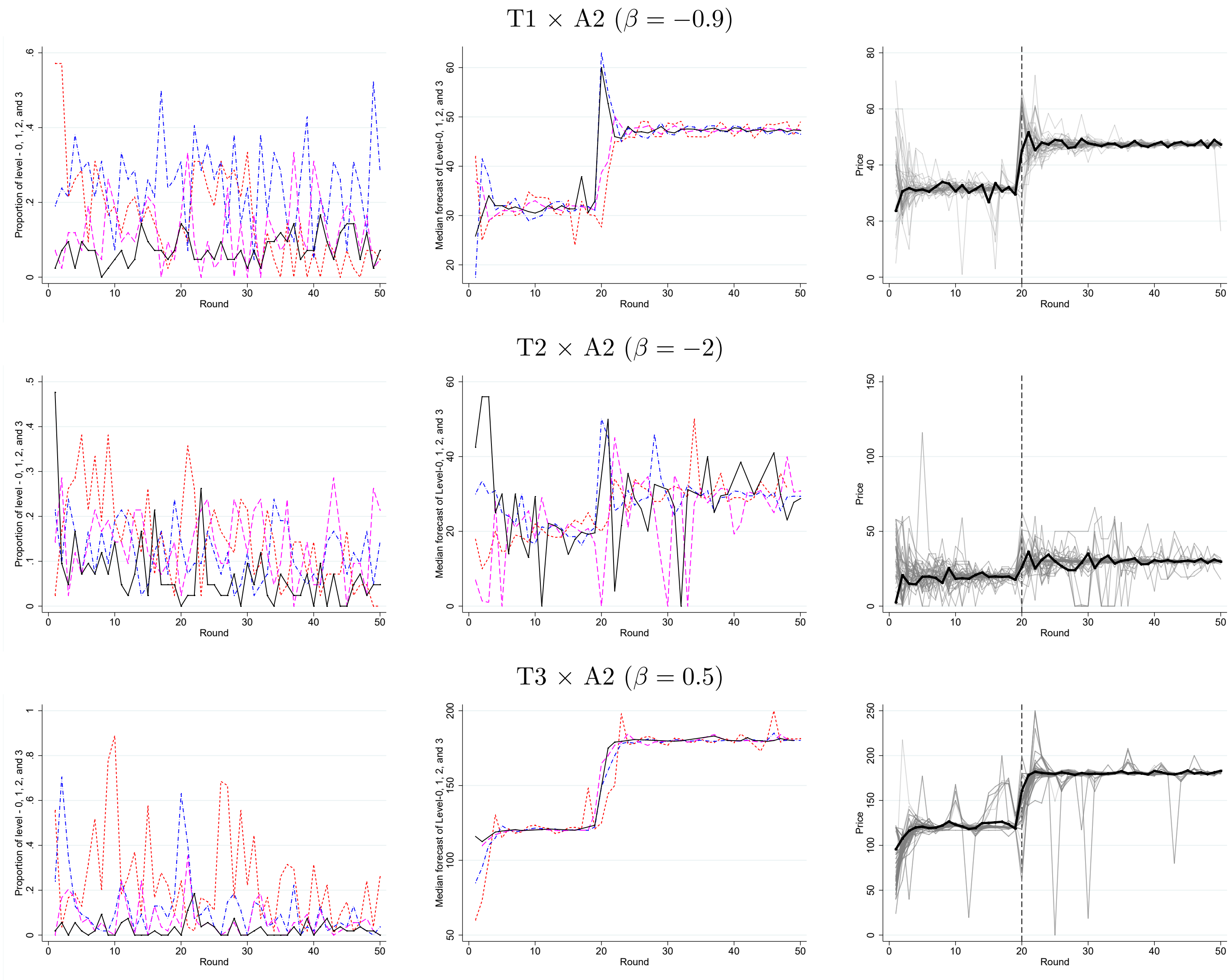
Table A1: Classification of predictions using counterfactual forecast rules

	All observations			Announcement periods only		
Treatment	RMSE	RMedSE	70^{th} Pctl	RMSE	RMedSE	70^{th} Pctl
T1 x A1	14.92	0.31	0.57	73.61	1.00	2.35
T1 x A2	10.53	0.36	0.64	7.78	1.49	4.90
T1 x A3	9.73	0.37	0.78	28.97	1.70	4.35
T2 x A1	9.59	0.30	0.73	7.29	4.00	5.00
T2 x A2	21.78	0.36	1.05	7.34	3.00	5.56
T2 x A3	3.97	0.50	1.32	5.39	3.00	5.00
T3 x A1	22.53	0.50	1.00	12.04	2.00	9.07
T3 x A2	14.72	0.44	1.00	28.07	2.01	5.00

Notes: This table shows how well laboratory participants' forecasts can be classified using a counterfactual forecast. For each subject we construct Level-0, 1, 2, 3, and REE forecasts based on the observed market data available to participants at each point in time. We calculate the difference between this forecast and the observed forecast submitted by the participant. We classify the subject as Level-0, 1, 2, 3, or REE based on which comparison yields the lowest squared error. The table reports the root mean (RMSE), median (RMedSE), and 70^{th} percentile of the squared difference between the submitted forecast and the nearest counterfactual forecast. The 70^{th} percentile is shown because we were able to classify 70% of forecasts in announcement periods using a ±4.5 cutoff when the data is pooled.

[3] Inclusion of these outliers actually makes some of our results stronger. For example, with respect to the result reported in Table 5 in the main text, the forecast errors generated by some of these outliers move the results in favor of the unified model.

Figure A3: Comparing the unified model to experimental data



Notes: Survey participants' forecasts are classified as Level-0, 1, 2, 3, or consistent with the REE forecast by comparing to the model implied forecasts. The time path of observed ω_n for $n = 0, 1, 2, 3$ are distinguished by plot-style: *red dotted, blue dash-dot, magenta dash* and *black solid*, respectively. The corresponding median forecasts, $E_{t-1}y_t^k$, of the participants use the same style format. The final column shows average market prices observed (solid black) laid over all individual forecasts. We omitted some outliers from the the final column of figures, which are shown in Figure A2 for clarity. The omitted forecasts are included in the calculations in the first two columns.

Figure A3 provides the same data breakdown for A2 treatments that we provided for A3 treatments in Figure 6 in the main text. We find similar results here. We identify heterogeneous forecasts that display level-k depths of reasoning in announcement rounds with mdiean individual and mean market dynamics closely matching what was predicted in Section 3.4 in the main text.

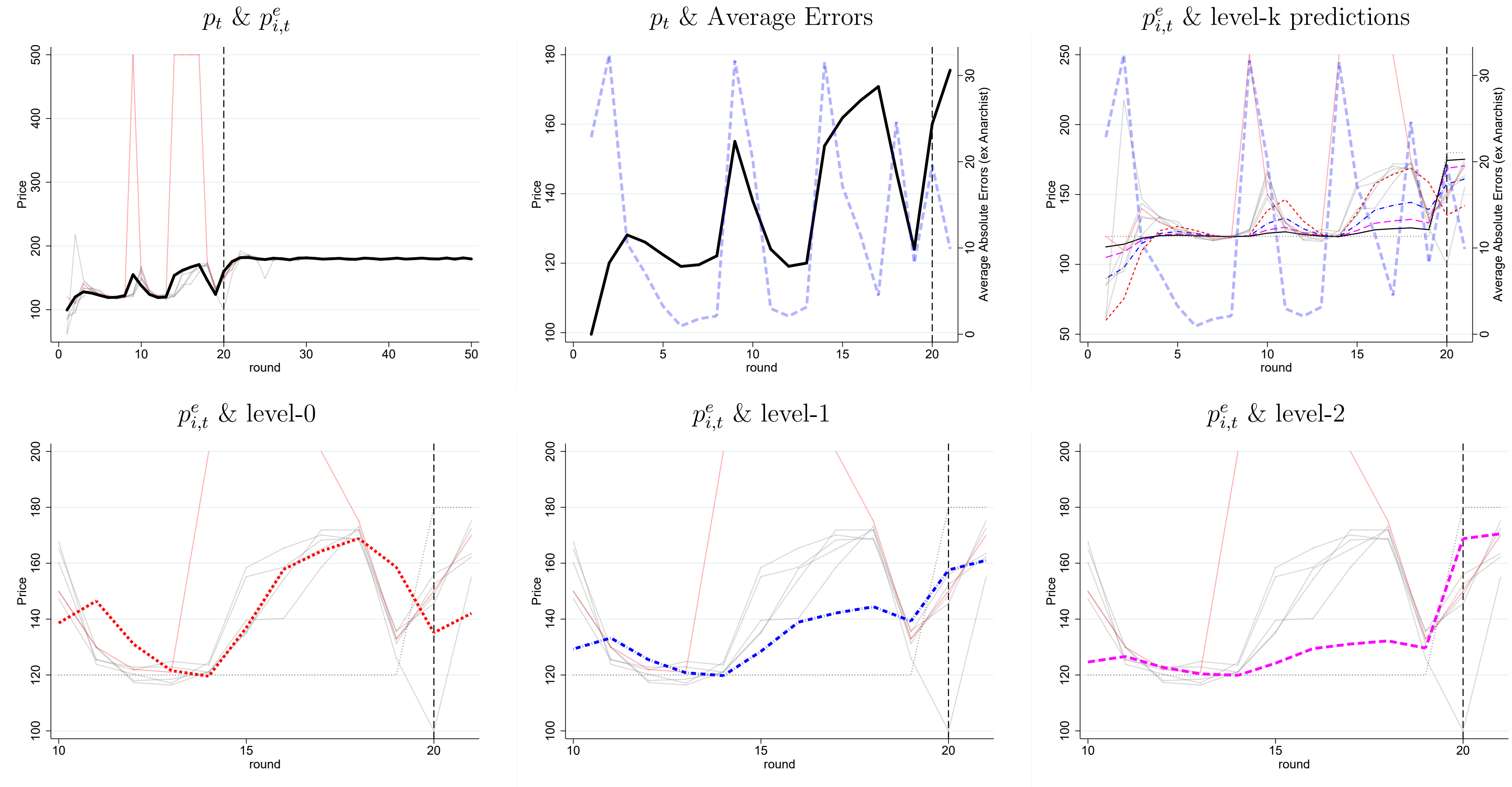
There are some additional points of interest in Figure A3 worthy of comment. In the T3 treatment we observed one market that had a significant departures in price from the REE after a period of convergence to the REE. You can see the individual forecasts in the third graph on the right of the last row. There is a group of individual forecasts that rise for many periods prior to the announcement in period 20. This market is what causes the spike in the median level-0 forecast that can be seen in the middle figure on the bottom row of Figure A3. The cause of this divergence is an anarchist player. This player's actions provide a nice case study for the unified model. For the five players who are attempting to play the game normally, the market has both large unobserved shocks and announced shocks.

A2.1 An Anarchist Anecdote

Figure A4 and A5 provide some detail on this anarchist's market. The first graph in the top left of Figure A4 shows the market price and the individual forecasts of the market participants. The anarchist is shown in red. The market converged to the REE by period 7. The anarchist then decided in period 9 to enters a price of 500, which was the largest price that the program would allow. The next figure shows the result. The price increased and a significant forecast error was realized by all other market participants. The anarchist struck again in round 14 and this time repeatedly enter a price of 500 for four consecutive rounds (ending in round 17). As before, there is a significant forecast error realized by all other players in the period the anarchists defects. However, the players quickly adapt to this unexplained rise in the price and the average forecast error falls over the next four periods. Importantly, we see all players switching to a forecast that lines up well with an adaptive forecast, consistent with the assumptions of the unified model. When the anarchist switches strategy in round 18, another large forecast error is generated, which causes yet another clear change in the strategy choices among the other participants.

The final figure in the top row of A4 layers onto the individual expectations the implied level-0, 1, 2, and 3 forecasts using our standard assumptions from Section 5. The bottom row of figures in A4 zooms in on the period of interest and plots the implied path of a single level-k forecast on each graph for clarity. It is immediately apparent that each large forecast error generates a shift in behavior by the non-anarchist players. Each shift in behavior is well-captured by one of the level-k deductions.

Figure A4: An Anarchist Anecdote



Notes: These figures show data from a T3×A2 treatment experimental market, where one player decided to actively sabotage the market. The anarchist's forecasts are shown in red. The time path of the implied level-0, 1, 2, and 3 forecasts are distinguished by plot-style: *red dotted, blue dash-dot, magenta dash* and *black solid*, respectively. The REE forecast is *black dotted*. The market price is the *solid thick black line.*

To see this, start by looking at period 10. Recall that there is no information that the participants have to suggest why the price suddenly moved in period 9. All participants trend follow in period 10 and revise their forecasts up. But the anarchist reverses course and provides a reasonable forecast in period 10, this generates another sizable forecast error. For period 11, the other participants switch strategies again. They appear to revise up their depth of reasoning and predict that market price will again fall. Both level-1 and level-2 predictions, which are based on the average price for rounds 9 and 10, explain nearly all the variation in forecasts chosen in this period. This switch by participants to a higher level strategy in period 11 generates a low forecast error and the subjects appear to maintain these strategies in the subsequent periods leading the market to converge.

When the anarchists strikes again in period 14, the remaining participants are quick to revise their depth of reasoning down to level-0. Forecast errors fall when switching to this strategy so they maintain the level-0 strategy. When the anarchists stop choosing 500 and reverts to choosing a normal strategy, another large forecast error is realized by the other market participants. This leads to a change in strategy in the next round. The revised strategies observed in the next round all sit on, or between, the implied level-1 and level-2 strategies (see bottom row of plots in Figure A4).

The chaos of this market is distinct from most other markets we observed. This raises the question of what the participants will do in an announcement round after the market has been so unpredictable. It appears that they mostly respond in accordance with the unified model. Five out the six forecasts for the announcement round sit between the level-1 forecast using our standard definition and a level-1 forecast where the level-0 assumptions is $p = 120$, which is the steady state price prior to the announced change.

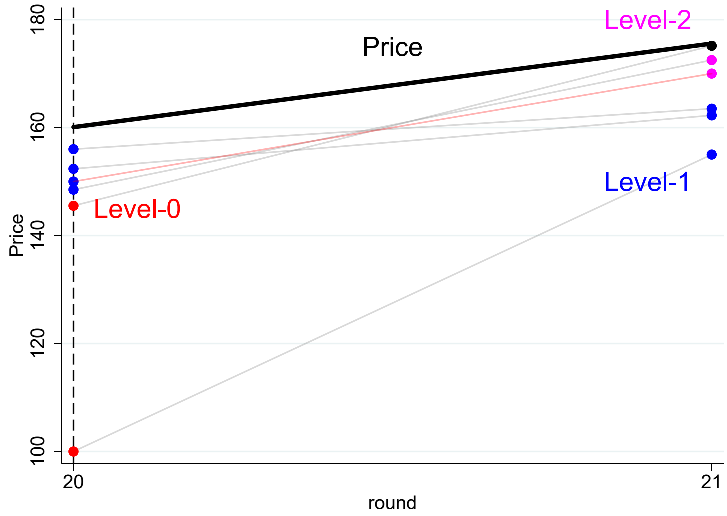
Figure A5 zooms in even further on just rounds 20 and 21 and classifies the individual forecasts types using the method described in Section 5 in the main text. Between the two rounds of play, those subjects whose forecasts were closest to the actual price, i.e. experienced the smallest errors, stick with the level-1 forecast. Those subjects who experience larger errors clearly revise up their depth of reasoning, where a revision to level-2 corresponds to what would have been the best forecast to play in round 20 given what occurred. This behavior is consistent with the assumptions that underlie the replicator dynamic's reflective process that we assume for the unified model.

A3 Robustness: Level-0 forecast definition

To classify the types of forecasting strategies that participants use, we must assume a shared level-0 forecast. Our baseline assumption is that level-0 is a two-round moving average of past prices. To demonstrate that our results are robust to this assumption, we conduct two exercises. First, we replicate the results in Table 4 and 5 of the main text using a four round moving average as the shared level-0 forecast. Second, we study how overall classification of types and of the level-0 type changes when we assume last periods price as the level-0 forecast, a two-period moving average, a four-period moving average, or three different cosntant gain specifications.

Table A4 replicates Table 4. The number of people we classify as level-k

Figure A5: An Anarchist Anecdote Announcement Round



Notes: Individual classified price forecasts in a T3×A2 treatment. The classifications are made by comparing the forecasts to different implied level-k forecasts. The closest implied forecast type determines the classification (see Section 5).

reasoners increases slightly under this definition overall. The regression estimates are mostly unchanged. We retain statistical significance for the hypothesis test conducted on announcement rounds with a comparable F-stat obtained to the original specification.

Table A3 replicates Table 5 for the four-round average level-0 assumption. The results are slightly stronger on all categories relative the previous definition.

Table A4 shows the classification results for the ±3 cut off for different level-0 assumptions. In general, the proportion of subjects that we classify as level-k forecasters of any type increases as we consider level-0 forecasts with longer averages or weighted averages of past observed prices.

A4 Oscillating deductions with strategic substitutes

García-Schmidt and Woodford (2019) and Angeletos and Sastry (2021) both raise the possibility that level-k reasoning may be implausible when there is strategic substitutability, i.e. negative feedback to expectations. In such an environment level-k deductions imply an oscillatory pattern in which each deduction takes expectations from one side of the perfect foresight equilibrium to the other. For example, this is clearly visible in the simulation of the unified model in Figure 4 in the main text. The third column of the figure shows the predicted level-1, 2, and 3 forecasts, which are on opposite sides of $\bar{y}$. It is argued that it is more plausible that agents would not contemplate such oscillations and would instead think about monotonic convergence from either above or below the perfect foresight equilibrium. Our experimental setting offers an environment that can shed light on whether people are willing to entertain oscillating deductions at least in this simple environment. We of course cannot observe the thoughts of subjects in each period, but we can observe their actions over time and in announcement rounds.

Announcement rounds in particular offer an interesting insight into people's thinking, especially when there is significant negative feedback as in the T2 treat-

Table A2: Classifying participant's forecasts as Level-k - Robustness Check

Within ±3 of Level-k in announcement rounds	1	20/50	45
Total Classified	47.3%	65.8%	66.0%
	[33.9% , 57.0%]	[50.6% , 73.6%]	[49.4% , 69.9%]
Level-0	14.8%	7.8%	5.1%
	[11.0% , 15.1%]	[4.31% , 9.48%]	[4.49% , 6.41%]
Level-1	7.3%	25.0%	14.1%
	[6.45% , 8.60%]	[19.3% , 27.6%]	[14.1% , 14.1%]
Level-2	6.5%	5.2%	3.8%
	[1.89% , 6.45%]	[4.31% , 5.75%]	[1.92% , 3.85%]
Level-3	3.2%	3.2%	4.5%
	[1.07% , 1.13%]	[2.58% , 3.74%]	[3.21% , 5.13%]
REE	15.6%	24.7%	38.5%
	[13.4% , 15.6%]	[20.1% , 27.0%]	[25.6% , 40.4%]
N	372	348	156

Hypothesis tests of deliberation time regressions	
H_0 : Level-0 - Level-3 = 0	F(1, 61) =0.75
H_0 : (Level-0 x Ann) - (Level-3 x Ann) = 0	F(1, 61) =**4.33**

Differences in deliberation time (seconds)		
Variable	[1]	[2]
Level-0	**-9.73**	**-1.97**
	(1.113)	(0.672)
Level-1	**-8.61**	-0.16
	(1.144)	(0.629)
Level-2	**-5.50**	-0.24
	(1.312)	(0.991)
Level-3	**-6.67**	-0.33
	(1.332)	(1.054)
Level-0 x Ann	**47.52**	3.58
	(8.623)	(6.065)
Level-1 x Ann	**45.47**	**11.90**
	(4.881)	(4.741)
Level-2 x Ann	**8.61**	11.08
	(9.058)	(8.546)
Level-3 x Ann	**63.82**	**23.02**
	(11.92)	(8.260)
Cons	**41.16**	**112.52**
	(0.526)	(4.227)
Individual FE	yes	yes
Round FE	no	yes
R-squared	0.030	0.253
N	18,367	18,367

Notes: The top left panel reports the proportion of participant's forecasts that fall within ±3 of a Level-k forecast. Proportions for cutoffs of ±1.5 and ±4.5 are shown in brackets. The right panel reports the regression results of identified Level-k individual's deliberation time in all periods and in announcement periods. Standard errors are clustered at the market level and reported in parenthesis below the point estimates. Bolded values indicate statistical significance at the ten percent level. The bottom left panel reports the hypothesis tests for the equality of regression coefficients for regression specification (2). We pool A1 (round 50 announcement) and A2 (round 20 announcement) results because both experiments feature a single and identical announcement.

ments ($\beta = -2$). The level-k deductions quickly push people towards a forecast of zero or γ in this case, which are prices that most players have never observed in any of the previous rounds that they have played. In fact, in most markets the price is on average at the REE for many periods before the announcement, which makes picking some other price, especially one that is far from the new equilibrium and on the opposite side it from the current price, a clear signal that people are contemplating oscillating deductions and acting on them. In Figure 6 and Figure 7 of the main text, we show clearly that people do indeed make predictions consistent with level-k deductions on either side of the perfect foresight equilibrium in an announcement rounds when $\beta < 0$. The T2 treatments arguably provide the strongest evidence of level-k behavior out of all or our treatments.

Another way to explore this hypothesis is to observe whether individual forecasts converge monotonically over time to the REE or whether they oscillate above and below over the REE price over time when $\beta < 0$. On this question we find heterogeneity across individual markets. The most common behavior is a clear willingness to contemplate oscillating deductions, but we also find a smaller number of episodes of clear monotonic behavior as well.

Figures A6 and A7 show two examples of markets with clear oscillating behavior and Figures A8 shows one market with more monotonic behavior. The top row of plots in each figure summarizes the market dynamics in each case with the first plot showing the market price and all of the individual forecasts, the second plot showing the individual forecasts plotted against the level-0,1,2, and

Table A3: Revisions and loss - Robustness Check

Treatment	Proportion of changers Between rounds 20 & 21		Ave. abs. prediction error Round 20			Ave. deliberation time (sec) Round 21		
	Revise opt.	No Change	Change	No change	Difference	Change	No change	Difference
T1 x A2 and A3	**0.62**	0.37	18.43	7.08	**11.35**	58.3	47.8	10.48
	[5.75]	(31/84)			[6.11]			[1.27]
T2 x A2 and A3	**0.48**	0.51	24.95	13.16	**11.79**	62.7	56.5	6.16
	[3.20]	(46/90)			[4.41]			[0.69]
T3 x A2 and A3	**0.56**	0.25	26.83	14.84	**11.99**	36.5	35.8	0.73
	[6.24]	(30/119)			[3.31]			[0.12]
	Between rounds 45 & 46		Round 45			Round 46		
T1 x A3	**0.73**	0.64	26.4	5.34	**21.06**	43.3	30.0	**13.3**
	[3.97]	(27/42)			[6.39]			[1.35]
T2 x A3	0.30	0.58	17.42	11.77	**5.65**	36.1	27.8	**8.23**
	[0.47]	(28/48)			[1.48]			[1.32]
T3 x A3	**0.52**	0.15	31.18	21.58	**9.60**	25.4	18.2	**7.16**
	[4.25]	(10/66)			[1.56]			[2.21]

Notes: "Revise opt." is the proportion of people who, conditioning on changing their strategy in period 21(46), changed their strategy to the best counterfactual strategy out of level-0, 1, 2, 3, or the REE in their market, where best is defined as what forecast would have been best in round 20(45). Z-scores for the test of the null hypothesis that subjects switched to one of the five strategies at random are reported in brackets. The next column reports the proportion of participants who we classify as not changing their strategy either between rounds 20 and 21 or between rounds 45 and 46 following announcements in either round 20 or 45, respectively. Counts appear in parentheses below. The remaining columns report the difference in average absolute prediction errors and average deliberation time for subjects classified as changing versus not changing with two-sample t-test statistics reported in brackets. Bolded values represent statistical significance at the ten percent level.

Table A4: Classifying participant's forecasts as Level-k - Robustness Check

	Moving averages			Constant Gain		
Round 20/50	1-period	2-period*	4-period	$\phi = 0.4$	$\phi = 0.3$	$\phi = 0.2$
Level - k	63.8%	64.4%	65.8%	65.2%	66.1%	66.1%
Level - 0	6.6%	6.6%	7.8%	6.9%	7.2%	7.2%

Notes: *Assumption used for level-k classification in the main text. The table reports the proportion of participant's forecasts that fall within ±3 of a Level-k forecast in all treatments with an announcement in period 20 or 50. The level-k forecasts are based on the level-0 assumption denoted in the table.

REE forecasts, and the third plot showing the classification of each type along with just the level-0 forecast.[4] The latter two plots are zoomed in around the announcement period. The bottom six plots of each figure show each individual's forecasts from each market classified by level-k type period-by-period compared to just the level-0 forecast. These plots illustrate the evolution of an individual's forecasts over time relative to the principle reference point for level-k deductions: the level-0 forecast.

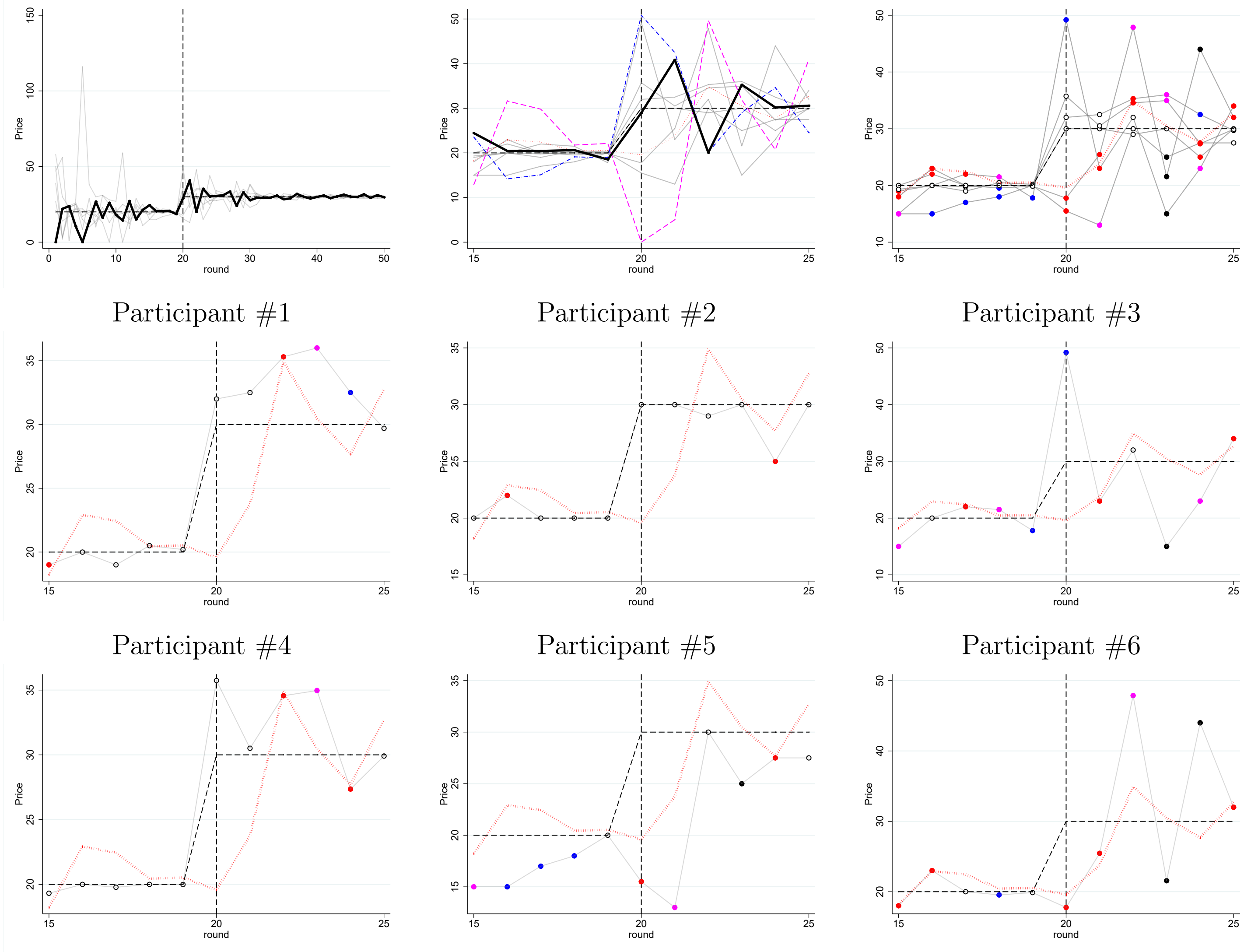
Figures A6, A7, and A8 clearly suggest a willingness by some individuals to oscillate their predictions above and below both the level-0 forecast and the REE price. The oscillations occur despite the experience of the price not oscillating for many periods prior to the announcement. This experience of tranquility combined with how close many of the forecasts are to level-k deductions is at least

[4] We do not plot the level-3 forecast in the middle figure because it makes the graph harder to interpret by requiring a larger scale of the y-axis. For the markets we show, no one chooses it in the announcement round. This of course is not true in general. We observe people choosing exactly level-3 deductions in some markets as can be seen in Figure 6 in the main text.

suggestive evidence that people contemplated oscillations consistent with classic level-k reasoning. Moreover, they took action with money at stake consistent with such deductions.

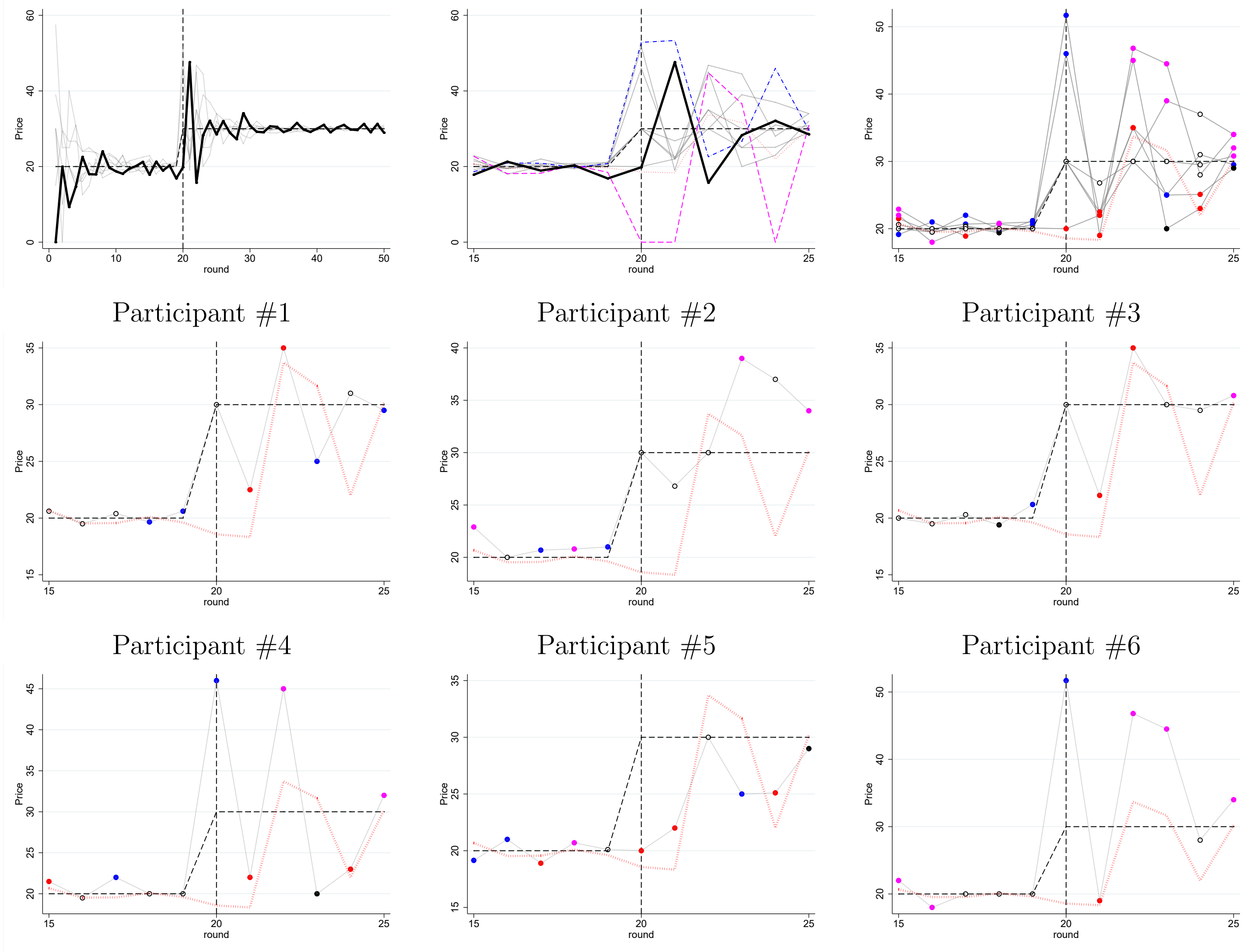
To further illustrate point, Figures A9 show the same type of analysis applied to a T3 treatment where $\beta = 0.5$. Level-k deductions do not imply oscillations in this case and indeed none are observed. Individual forecasts conform nicely to level-k deductions based on our proposed level-0 forecast. This suggests that people do not abandon level-k deductions in environments with strategy substitutability. Level-k deductions describes forecasting behavior in our experiment when strategic actions are both compliments and substitutes.

Figure A6: Example 1: Individual forecasts from experimental market with treatment T2 ($\beta = -2$)



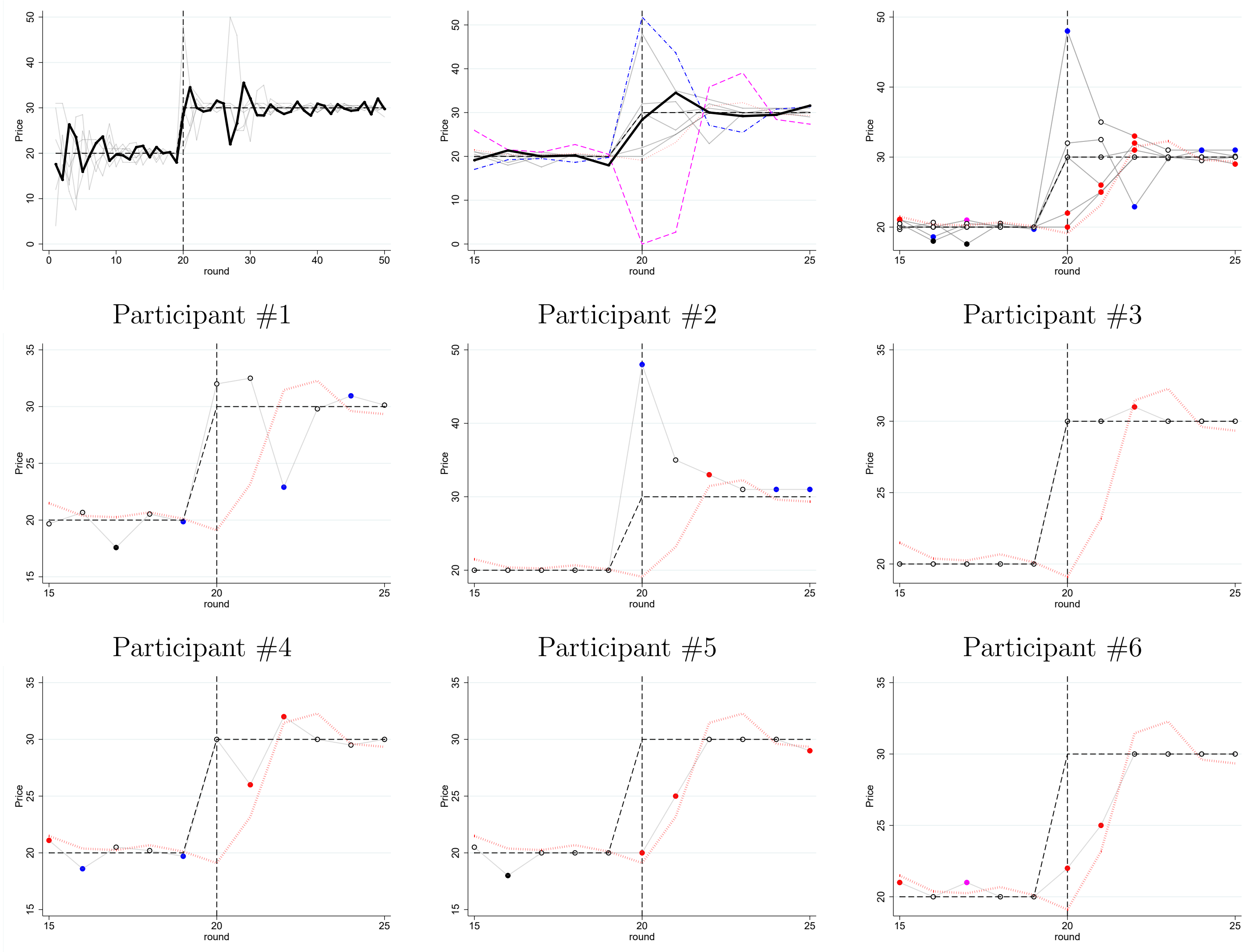
Notes: The first plots shows all individual forecasts and the market price from a single market. The second plot shows the model implied level-k forecasts and the data in a window around the announcement. The remaining figures classify each of the forecasts as a level-k type, which is indicated by the color of the dot. Forecasts that are classified as level-0 are shown in red, level-1 in blue, level-2 in magenta, level-3 in black, and REE as a black circle with a white interior. The dotted red line shows the path of the level-0 forecast from which all level-k deductions are derived. The perfect foresight equilibrium is indicated by the dashed line.

Figure A7: Example 2: Individual forecasts from experimental market with treatment T2 ($\beta = -2$)



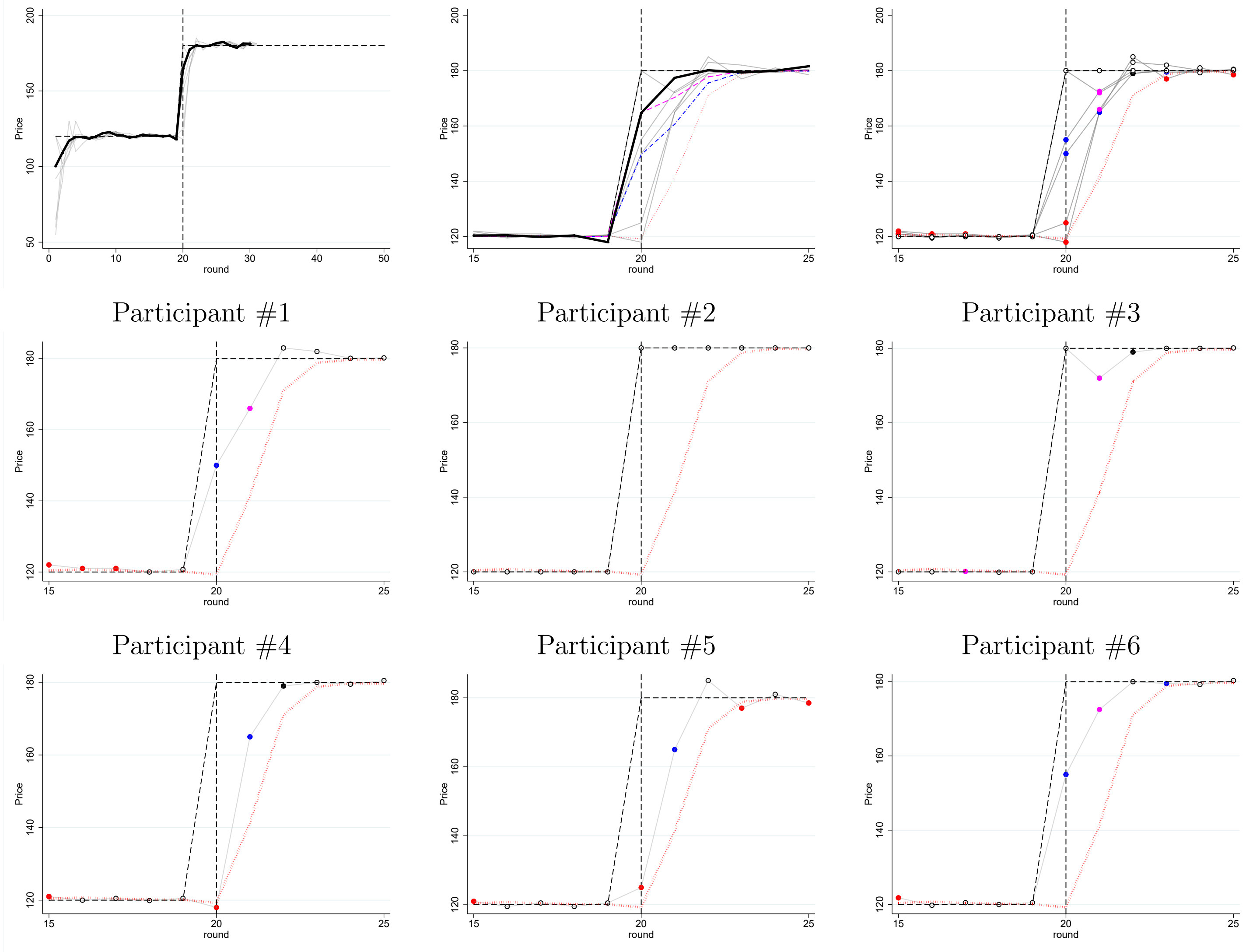
Notes: The first plots shows all individual forecasts and the market price from a single market. The second plot shows the model implied level-k forecasts and the data in a window around the announcement. The remaining figures classify each of the forecasts as a level-k type, which is indicated by the color of the dot. Forecasts that are classified as level-0 are shown in red, level-1 in blue, level-2 in magenta, level-3 in black, and REE as a black circle with a white interior. The dotted red line shows the path of the level-0 forecast from which all level-k deductions are derived. The perfect foresight equilibrium is indicated by the dashed line.

Figure A8: Example 3: Individual forecasts from experimental market with treatment T2 ($\beta = -2$)



Notes: The first plots shows all individual forecasts and the market price from a single market. The second plot shows the model implied level-k forecasts and the data in a window around the announcement. The remaining figures classify each of the forecasts as a level-k type, which is indicated by the color of the dot. Forecasts that are classified as level-0 are shown in red, level-1 in blue, level-2 in magenta, level-3 in black, and REE as a black circle with a white interior. The dotted red line shows the path of the level-0 forecast from which all level-k deductions are derived. The perfect foresight equilibrium is indicated by the dashed line.

Figure A9: Example 4: Individual forecasts from experimental market with treatment T3 ($\beta = 0.5$)



Notes: The first plots shows all individual forecasts and the market price from a single market. The second plot shows the model implied level-k forecasts and the data in a window around the announcement. The remaining figures classify each of the forecasts as a level-k type, which is indicated by the color of the dot. Forecasts that are classified as level-0 are shown in red, level-1 in blue, level-2 in magenta, level-3 in black, and REE as a black circle with a white interior. The dotted red line shows the path of the level-0 forecast from which all level-k deductions are derived. The perfect foresight equilibrium is indicated by the dashed line.

A5 Exit Survey Results

After the experimented ended, subjects completed an exit survey while they waited for their pay envelopes to be prepared. The survey questions aimed to assess what information they used to make their forecasts and what information they thought others used.

A5.1 Exit Survey Questions

1. Please rank the importance of each option below to the formation of your price forecast in each period:

   a. The history of market prices

   b. The market equations

   c. The history of my own price forecasts

   d. The history of my own forecasts errors

   e. My expectation about the average price forecast in the period

2. Please rank the importance of each option below to the formation of your price forecast **following the announcements**:

   a. The history of market prices

   b. The market equations

   c. The history of my own price forecasts

   d. The history of my own forecasts errors

   e. My expectation about the average price forecast in the period

3. Which of the following statements best describes your thinking before making each forecast?

   a. I looked at the past prices and made my best guess based on their recent movements. I never used the equations.

   b. I made a guess about what the average forecast might be based on past prices and then used the equations to determine my own forecast using that guess.

   c. I made a guess about what the average forecast might be and used the equation to work out the price only when I did a poor job of forecasting in the previous round. Otherwise, I just looked at past prices and made my best guess.

   d. I made a guess about what the average forecast might be and used the equation to work out the price only when there was an announced change in the market. Otherwise, I just looked at past prices and made my best guess.

4. Please rank the importance of each option below to **other participants**, which you believe they may have used to make their price forecasts:

   a. The history of market prices

   b. The market equations

   c. The history of their own price forecasts

   d. The history of their own forecasts errors

   e. Their expectation about the average price forecast in the period

5. Please rank the importance of each option below to **other participants**, which you believe they may have used to make their price forecasts **following the announcements:**

   a. The history of market prices

   b. The market equations

   c. The history of their own price forecasts

   d. The history of their own forecasts errors

   e. Their expectation about the average price forecast in the period

6. If you do not feel like the strategy you used was well-captured by the survey questions, then please use this box to explain your strategy

A5.2 Exit Survey Results

Survey questions (1), (2), (4), and (5) used a drop-down menu with options: "very important", "somewhat important", and "did not consider." Table A5 and Table A6 shows the cumulative importance of each factor where "very important" is assigned a zero, "somewhat important" is assigned a one, and "did not consider" a two. Therefore, the lower the value, the more important the information. Consistent with level-k reasoning, we find that on average subjects rated the equations and the forecast of the average expectation as more important to their own forecast than they believed it was to others. This is consistent with a belief that others are less sophisticated. We observe the results on the full sample and when restricting to only people who played a level-k forecast in the announcement periods with the $\pm$ 3 cutoff. The latter consistently rank the equations as important to them than they are to their perceived competitors, which is consistent with the level-k assumption that others players are perceived as less sophisticated.

Figure A10 shows the responses to question 3 separated by treatment. The most common response is (b), which is:

> *I made a guess about what the average forecast might be based on past prices and then used the equations to determine my own forecast using that guess.*

This response is consistent with level-1 behavior.

Table A5: Tabulated survey results for Q1 and Q4

All Responses										
	Past Prices		Equations		Forecast History		Forecast Errors		Exp. Ave. Price	
Treatment	Own	Others	Own	Others	Own	Others	Own	Others	Own	Others
T1×A1	10	9	33	32	24	21	30	32	19	21
T1×A2	11	11	32	30	33	19	38	31	17	19
T1×A3	11	11	31	30	40	31	45	32	21	23
T2×A1	13	8	25	23	31	25	36	30	16	19
T2×A2	13	10	28	28	47	24	50	42	18	30
T2×A3	23	18	36	41	52	35	47	43	25	41
T3×A2	7	7	42	37	50	33	44	42	20	29
T3×A3	16	13	47	59	67	52	63	66	30	33
All	104	87	274	280	344	240	353	318	166	215
Difference	17		-6		104		35		-49	
Info is (__) to me	(less important)		(more important)		(less important)		(less important)		(more important)	
Responses from those identified as level-k in announcement rounds with ±3 cutoff										
	Past Prices		Equations		Forecast History		Forecast Errors		Exp. Ave. Price	
Treatment	Own	Others	Own	Others	Own	Others	Own	Others	Own	Others
T1×A1	6	6	24	22	13	10	17	18	10	13
T1×A2	5	7	10	14	18	12	21	15	7	9
T1×A3	4	3	3	7	8	9	8	8	3	5
T2×A1	7	3	8	11	14	9	13	9	8	9
T2×A2	4	3	2	7	18	12	23	20	6	14
T2×A3	3	4	1	6	6	4	5	2	2	10
T3×A2	0	0	5	6	14	10	14	9	6	9
T3×A3	6	4	10	14	12	12	11	14	4	8
All	35	30	63	87	103	78	112	95	46	77
Difference	5		-24		25		17		-31	
Info is (__) to me	(less important)		(more important)		(less important)		(less important)		(more important)	

Notes: Participants rated each piece of information denoted in the top line as "very important", "somewhat important", or "did not consider" when making their "own" forecasts and what they believed was important to "others". The categories are assigned the following values and summed: "very important" is a assigned a zero, "somewhat important" a one, and "did not consider" as two. Lower totals indicate that the piece of information is more important to a person's decision.

A6 Quantitative evaluation: additional results

We fit the model to the experimental data at the market level. Table 6 in the main text averages over the individual market outcomes from the same treatments. Table A7 shows the underlying data from each market.

Each model that features heterogeneous types is initialized to the first realized price and to the distribution of level-k types observed in period one for each market. Afterwards, the model makes predictions based solely on the evolution of price, adaptive learning, or the replicator, depending on which model is used. The learning model starts initial beliefs at the average of the individual forecasts in period one. After period one it updates according to the evolution of data implied by the model and beliefs for the chosen gain. The simulated data is compared to experimental data and the mean squared error is calculated.

Each model is optimized individually by searching over a grid of gains $\phi \in [0,1]$, or replicator parameters $\alpha \in [0,2]$, or both in the case of the unified model. The optimal coefficients are shown in Table A8. Both the replicator and adaptive learning are required to best fit the data in T1 and T2 treatments. In many of the T3 treatments, however, naive expectations and fixed level-k reasoning is chosen as the best model. This reflects the fact that many markets coverge very

Table A6: Tabulated survey results for Q2 and Q5

Treatment	Past Prices Own	Past Prices Others	Equations Own	Equations Others	Forecast History Own	Forecast History Others	Forecast Errors Own	Forecast Errors Others	Exp. Ave. Price Own	Exp. Ave. Price Others
All Responses										
T1×A1	13	12	30	29	24	24	29	39	23	18
T1×A2	19	19	24	19	42	35	36	32	23	21
T1×A3	19	19	28	24	40	29	44	38	24	23
T2×A1	11	21	18	23	33	27	40	34	21	20
T2×A2	21	19	21	20	44	24	49	41	20	24
T2×A3	31	35	30	36	52	49	54	45	32	35
T3×A2	25	27	31	25	51	39	54	45	26	28
T3×A3	40	44	38	38	75	64	73	76	28	30
All	179	196	220	214	361	291	379	350	197	199
Difference	-17		6		70		29		-2	
Info is (__) to me	(more important)		(less important)		(less important)		(less important)		(more important)	
Responses from those identified as level-k in announcement rounds with ±3 cutoff										
T1×A1	7	7	22	21	12	15	18	23	13	10
T1×A2	11	9	6	4	22	19	16	19	9	12
T1×A3	4	3	4	6	8	8	8	10	5	5
T2×A1	7	7	7	12	16	8	16	13	9	11
T2×A2	9	8	1	4	19	10	22	18	14	8
T2×A3	4	4	1	5	5	4	4	1	10	7
T3×A2	1	4	7	9	15	9	18	10	9	7
T3×A3	11	11	8	7	19	12	15	17	8	7
All	54	53	56	68	116	85	117	111	77	67
Difference	1		-12		31		6		10	
Info is (__) to me	(less important)		(more important)		(less important)		(less important)		(less important)	

Notes: Participants rated each piece of information denoted in the top line as "very important", "somewhat important", or "did not consider" when making their "own" forecasts and what they believed was important to "others". The categories are assigned the following values and summed: "very important" is a assigned a zero, "somewhat important" a one, and "did not consider" as two. Lower totals indicate that the piece of information is more important to a person's decision.

quickly to steady state, but not as quickly as RE implies. This is also reflected in the results for the adaptive learning case were a naive model is found to best fit the data for all markets. In subsequent exploration, which is not shown here, we have found that a $\phi > 1$ plus level-k reasoning is preferred. That is consistent with a trend following behavior similar to what many other positive feedback experiments have found.

A6.1 Further derivations for the NK model

For the NK model (14), inflation is determined as follows:

$$\pi_t = \gamma_t + \beta_t \sum_{k \geq 0} \omega_t(k) E_t^k \pi_{t+1}, \tag{A14}$$

where $E_t^k \pi_{t+1}$ is the period t forecast of π_{t+1} made by a k-level agent. An expression for $E_t^k \pi_{t+1}$ can be derived using backward induction:

$$\begin{aligned} E_t^k \pi_{t+1} &= \gamma_{t+1} + \beta_{t+1} E_t^{k-1} \pi_{t+2} \\ &\vdots \\ E_t^1 \pi_{t+k} &= \gamma_{t+k} + \beta_{t+k} a_t \end{aligned}$$

Table A7: MSE between experimental data and competing models

Treatment	REE	Unified Model		Fixed Level-k		Replicator only		Adaptive learning	
T1 × A3	MSE	MSE	Rel. REE	MSE	Rel. REE	MSE	Rel. REE	MSE	Rel. REE
Market 1	6.67	3.88	0.58	11.59	1.74	4.29	0.64	20.77	3.11
Market 2	8.66	4.17	0.48	13.34	1.54	4.17	0.48	27.38	3.16
Market 3	24.18	22.34	0.92	25.07	1.04	22.34	0.92	39.76	1.64
Market 4	14.47	3.01	0.21	3.87	0.27	3.51	0.24	29.48	2.04
Market 5	14.66	2.24	0.15	12.81	0.87	13.96	0.95	7.85	0.54
Market 6	18.20	4.11	0.23	16.94	0.93	17.74	0.97	17.23	0.95
Market 7	5.22	1.89	0.36	2.94	0.56	2.57	0.49	13.32	2.55
Average	13.15	5.95	0.45	12.37	0.94	9.80	0.74	22.26	1.69
T2 × A3									
Market 1	66.42	57.92	0.87	126.77	1.91	76.33	1.15	86.21	1.30
Market 2	25.31	20.44	0.81	154.20	6.09	34.67	1.37	34.06	1.35
Market 3	58.01	76.90	1.33	873.05	15.05	94.48	1.63	77.46	1.34
Market 4	48.70	40.98	0.84	779.52	16.01	75.16	1.54	65.73	1.35
Market 5	23.36	37.13	1.59	80.20	3.43	44.65	1.91	42.05	1.80
Market 6	44.84	51.04	1.14	569.37	12.70	69.59	1.55	68.70	1.53
Market 7	67.28	52.25	0.78	671.08	9.98	75.55	1.12	47.06	0.70
Market 8	80.64	50.42	0.63	127.50	1.58	97.45	1.21	85.84	1.06
Average	51.82	48.38	0.93	422.71	8.16	70.98	1.37	63.39	1.22
T3 × A3									
Market 1	22.49	1.80	0.08	1.80	0.08	35.16	1.56	36.23	1.61
Market 2	31.16	14.70	0.47	16.33	0.52	46.91	1.51	39.09	1.25
Market 3	37.48	17.64	0.47	17.64	0.47	42.97	1.15	38.23	1.02
Market 4	31.37	13.70	0.44	13.70	0.44	31.78	1.01	48.27	1.54
Market 5	12.52	4.34	0.35	4.34	0.35	28.54	2.28	48.90	3.91
Market 6	28.67	33.11	1.15	35.38	1.23	60.66	2.12	75.75	2.64
Market 7	45.41	23.82	0.52	27.08	0.60	61.14	1.35	46.89	1.03
Market 8	44.70	19.56	0.44	22.89	0.51	50.28	1.12	48.38	1.08
Market 9	92.75	76.53	0.83	76.53	0.83	104.95	1.13	106.02	1.14
Market 10	31.71	3.46	0.11	3.46	0.11	40.40	1.27	28.27	0.89
Market 11	30.64	9.45	0.31	9.45	0.31	41.05	1.34	39.53	1.29
Average	37.17	19.83	0.53	20.78	0.56	49.44	1.33	50.51	1.36

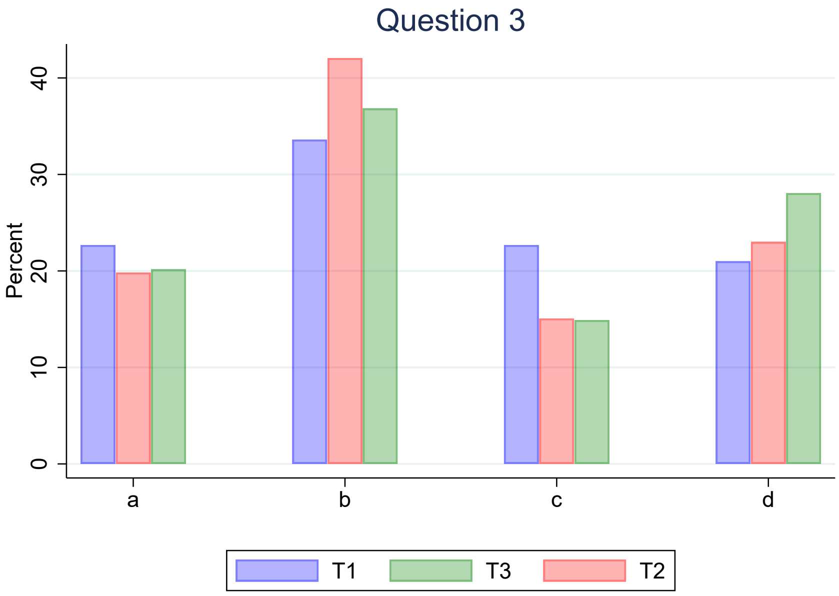
Notes: Mean square error (MSE) of five simulated models of aggregate price dynamics compared to experimental market price data. "Rel. REE" reports the MSE of the a model relative to REE MSE, i.e., Model MSE/REE MSE. Models are fit by doing a grid search over values $\alpha \in [0, 2]$ and $\phi \in [0, 1]$.

Table A8: Parameter estimates of competiting models

Treatment	Unified Model		Fixed Level-k		Replicator only		Adaptive learning
T1 × A3	α	ϕ	α	ϕ	α	ϕ	ϕ
Market 1	0.225	0.725	-	0.475	0.025	-	0.425
Market 2	0.015	0.000	-	0.550	0.015	-	0.500
Market 3	0.008	0.000	-	0.325	0.008	-	0.450
Market 4	0.200	0.775	-	1.000	0.005	-	0.475
Market 5	0.175	0.725	-	1.000	0.000	-	0.550
Market 6	0.150	0.750	-	1.000	0.000	-	0.575
Market 7	0.300	0.800	-	0.725	0.010	-	0.525
T2 × A3							
Market 1	0.005	0.100	-	0.000	0.005	-	0.325
Market 2	0.010	0.050	-	0.475	0.010	-	0.325
Market 3	0.010	0.050	-	0.125	0.010	-	0.400
Market 4	0.005	0.200	-	0.150	0.010	-	0.500
Market 5	0.015	0.025	-	0.175	0.010	-	0.300
Market 6	0.010	0.025	-	0.000	0.010	-	0.325
Market 7	0.005	0.175	-	0.150	0.010	-	0.525
Market 8	0.025	0.425	-	0.400	0.010	-	0.375
T3 × A3							
Market 1	0.000	1.000	-	1.000	0.175	-	1.000
Market 2	0.600	0.725	-	1.000	0.200	-	1.000
Market 3	0.000	1.000	-	1.000	0.175	-	1.000
Market 4	0.000	0.950	-	0.950	0.025	-	1.000
Market 5	0.000	1.000	-	1.000	0.075	-	1.000
Market 6	0.600	0.725	-	1.000	0.375	-	1.000
Market 7	0.600	0.725	-	1.000	0.350	-	1.000
Market 8	0.100	0.725	-	1.000	0.375	-	1.000
Market 9	0.000	1.000	-	1.000	0.200	-	1.000
Market 10	0.000	1.000	-	1.000	0.200	-	1.000
Market 11	0.000	1.000	-	1.000	0.225	-	1.000

Notes: Paremeter estimates of the competiting models. Models are fit by doing a grid search over values $\alpha \in [0, 2]$ and $\phi \in [0, 1]$. The Fixed level-k model assumes an adaptive level-0 forecast.

Figure A10: Exit survey question 3 responses



Notes: This figure shows the response to question 3 from exit survey separated by treatment type.

where $a_t = E_t^0 \pi_{t+m}$ for all m. Setting $\beta_t^k = \prod_{n=1}^k \beta_{t+n}$, it follows that

$$E_t^k \pi_{t+1} = \sum_{n=1}^{k} \beta_t^{n-1} \gamma_{t+n} + \beta_t^k a_t. \tag{A15}$$

Combining (A14) and (A15), the dynamics of the economy can be loosely written as $\pi_t = \gamma_t + \beta_t \pi_{t+1}^e$, however there is an important nuance. Under both rationality and unified learning, contemporaneous outcomes depend on (agents' perceptions of) the future path of model coefficients; and the entire path of coefficients is subject to change when policy announcements are made.

Let γ_{t+k}^t and β_{t+k}^t be the values of the coefficients in period $t+k$ as perceived by agents in period t. For example, we know that, in period T, agents think the the CB follows a perpetual Taylor rule: $\beta_{T+k}^T = \psi$ for all k. Also, in period $T+M_1$ the CB announces an interest rate peg for N periods, thus $\beta_{T+M_1+k}^{T+M_1} = \theta$ for $0 \leq k < N - M_1$. The dynamics of the economy are now given as

$$\pi_{t+k} = \gamma_{t+k}^t + \beta_{t+k}^t \pi_{t+k+1}^e \tag{A16}$$

where

$$T \leq t < T+M_1 \implies \gamma_{t+k}^t = \begin{cases} b\Delta v^* & \text{if } 0 \leq k < M-(t-T) \\ 0 & \text{else} \end{cases}$$

$$t \geq T+M_1 \implies \gamma_{t+k}^t = \begin{cases} b\Delta v^* - (\theta-1)\Delta i & \text{if } 0 \leq k < M_2 - (t-(T+M_1)) \\ -(\theta-1)\Delta i & \text{if } M_2 - (t-(T+M_1)) \leq k < N - M_2 - (t-(T+M)) \\ 0 & \text{else} \end{cases}$$

$$T \leq t < T+M_1 \implies \beta_{t+k}^t = \psi$$

$$t \geq T+M_1 \implies \beta_{t+k}^t = \begin{cases} \theta & \text{if } 0 \leq k < N - M_1 - (t-(T+M_1)) \\ \psi & \text{else} \end{cases}$$

References

Angeletos, George-Marios and Karthik A Sastry, "Managing Expectations: Instruments Versus Targets," *The Quarterly Journal of Economics*, 2021, *136* (4), 2467–2532.

García-Schmidt, Mariana and Michael Woodford, "Are low interest rates deflationary? A paradox of perfect-foresight analysis," *American Economic Review*, 2019, *109* (1), 86–120.

A7 Experiment materials

This section provides the instructions and tutorial information that were provided to laboratory subjects.

Negative feedback case

Computer based tutorial:

- What is your role?

  Your role is to act as an expert forecaster advising firms that produce widgets.

- What makes you an expert in this market?

  You will have access to information about the demand and supply of widgets to the market. You will also have a bit of training before making paid forecasts.

- What is a widget?

  Widgets are a perishable commodity like bananas or grapes. They are perishable in the sense that they can only be consumed in the period they are produced. They cannot be stored for consumption in future periods. The widgets that each firm produces are all the same and there are many firms in the market. Therefore, the individual firms do not set the price at which they sell their widgets but must sell widgets at the market price.

- Why do the firms need to forecast the price?

  A firm must commit to the number of widgets it will produce in the coming period before knowing the price. Therefore, the firms need to have a forecast of the price to know how many to produce.

- How am I paid?

  Your compensation for each forecast is based on the accuracy of the forecast. The payoff for each forecast is given by the following formula:

  $$\text{payment} = 0.50 - 0.03\,(p - \text{your price forecast})^2$$

  where p is the actual market price, and 0.50 and 0.03 are measured in cents. If your forecast is off by more than 4, you will receive \$0.00 for your forecast. Therefore, you will receive \$0.50 for a perfect forecast, where p=your price forecast, and potentially \$0.00 for a very poor forecast. You will be paid to make 50 forecasts in total.

  In addition, you will be paid a \$5 show-up fee for participating. You may quit the experiment at any time, for any reason, and retain this \$5 payment.

- The Demand for Widgets:

The total demand for widgets in a period is downward sloping. This means that the lower the price is the greater the demand for widgets. In precise terms, the demand is given by

$$q = A - Bp$$

where q is the quantity demanded, and p is the current price in the market. The equation for demand and the values for A and B will be given to you at the beginning of the experiment. The values may also change during the experiment. The equation, the values of A and B, and any changes to these values will be told to all participants at the same time.

- The Supply of Widgets:

  The firms in the market all face the same costs for producing widgets. The supply of widgets by each firm, therefore, only depends on their forecast for price next period. The total supply of widgets to the market depends on the average price forecast from all firms.

  The total amount of widgets supplied to the market by all firms is given by

$$q = D \times \text{average price forecast}$$

  where D is a positive number, which will be given to you and all other forecasters in the market at the start of the experiment. Just like with demand, D may change during the experiment and the changes will be announced.

- Prices and Expected Prices:

  Once all participants have chosen their expected price, the average expected price determines total supply. Since quantity demanded depends on price, equating supply and demand determines the price. Consequently the actual market price depends on average expected price. In fact there is a negative relationship between price and expected price. In other words, when the average forecast for the price is high, the actual price is low and vice versa.

- Why does this occur?

  It occurs because a high average expected price causes widget producers to increase their production of widgets. The increase in production results in more widgets supplied to the market. More supply of widgets means that the price of each widget will be lower. The opposite occurs when the average expected price is low. In this case, the widget producers will supply fewer widgets to the market, which results in a high price.

  By equating supply and demand,

$$A - Bp = D \times \text{average price forecast}$$

  we can arrive at the precise relationship for price and expected price

$$p = \frac{A}{B} - \frac{D}{B} \times \text{average price forecast}$$

  Note that expected price is negatively related to price. If expected price is high, then the actual price is low and vice versa

- A bit of randomness:

  Finally, like in real markets, we allow for the possibility that unforeseen and unpredictable things may happen that affect price. We add this to the game by adding a **small** amount of noise to price such that

$$p = \frac{A}{B} - \frac{D}{B} \times \text{average price forecast} + \textit{noise}.$$

The *noise* term is chosen at random in each period and is not predictable. Its value is not given to any participant in the market. The size of each realisation is **small**. The average value of the *noise* over the course of the experiment is zero and each realisation of it is independent from any other realization. In other words, the *noise* term may take a positive or a negative value in any given period, but overall, the size and number of positive and negative realisations will be approximately equal and cancel each other out over time.

Positive feedback case

Computer based tutorial:

- What is your role?

  Your role is to act as an expert forecaster advising firms that sell widgets.

- What makes you an expert in this market?

  You will have access to information about the demand and supply of widgets to the market. You will also have a bit of training before making paid forecasts.

- What is a widget?

  Widgets are a perishable commodity like bananas or grapes. They are perishable in the sense that they can only be consumed in the period they are produced. They cannot be stored for consumption in future periods. The widgets are all the same and there are many firms that sell in the market. Therefore, the individual firms do not set the price at which they sell their widgets but must sell widgets at the market price.

- Why do the firms need to forecast the price?

  Widgets are considered by many to be a luxury good, in part because they cannot be stored. In fact, when the price of widgets goes up, the demand for widgets tends to go up as well as many consider expensive widgets a status symbol. Therefore, how many widgets a firm should produce to meet demand depends on the expected price in the market that day. Each firm has an advisor like you that provides price forecasts. If the average price forecast is high, then firms will want to supply many widgets and the actual price will be high. If the average price forecast is low, then the firms will supply fewer widgets and the actual price will be low.

- How am I paid?

  Your compensation for each forecast is based on the accuracy of the forecast. The payoff for each forecast is given by the following formula:

$$\text{payment} = 0.50 - 0.03\,(p - \text{your price forecast})^2$$

  where p is the actual market price, and 0.50 and 0.03 are measured in cents. If your forecast is off by more than 4, you will receive \$0.00 for your forecast. Therefore, you will receive \$0.50 for a perfect forecast, where p=your price forecast, and potentially \$0.00 for a very poor forecast. You will be paid to make 50 forecasts in total.

  In addition, you will be paid a \$5 show-up fee for participating. You may quit the experiment at any time, for any reason, and retain this \$5 payment.

- The Demand for Widgets:

  The total demand for widgets in a period is upward sloping. This means that the higher the price, the greater the demand for widgets. In precise terms, the demand is given by

  $$q = A + Bp$$

  where q is the quantity demanded, and p is the current price in the market. The equation for demand and the values for A and B will be given to you at the beginning of the experiment. The values may also change during the experiment. The equation, the values of A and B, and any changes to these values will be told to all participants at the same time.

- The Supply of Widgets:

  The firms in the market all face the same costs for producing widgets. The supply of widgets by each firm, therefore, only depends on their advisor's forecast for price next period. The total supply of widgets to the market depends on the average price forecast from all firms.

  The total amount of widgets supplied to the market by all firms is given by

  $$q = C + D \times \text{average price forecast}$$

  where C and D are positive numbers, which will be given to you and all other forecasters in the market at the start of the experiment. Just like with demand, C and D may change during the experiment and the changes will be announced.

- Prices and Expected Prices:

  Once all advisors have chosen their expected price, the average expected price determines total supply. In each period, a central market-maker then sets the final price so that demand equals the quantity supplied. Consequently, the actual market price depends on the average expected price. In fact, there is a positive relationship between price and expected price. In other words, when the average forecast for the price is high, the actual price is high and vice versa.

- Why does this occur?

  It occurs because a high average expected price causes widget producers to increase their production of widgets. The higher the price, the higher the actual demand for widgets due the fact they are a status symbol. The opposite occurs when the average expected price is low. In this case, low prices will results in low demand as widgets appear to be less of a luxury good. By equating supply and demand,

  $$A + Bp = C + D \times \text{average price forecast}$$

  we can arrive at the precise relationship for the price and the expected price

  $$p = \frac{C - A}{B} + \frac{D}{B} \times \text{average price forecast}$$

  where we will assume that $C > A$. Note that the expected price is positively related to price. If the expected price is high, then the actual price is high and vice versa

- A bit of randomness:

Finally, like in real markets, we allow for the possibility that unforeseen and unpredictable things may happen that affect price. We add this to the game by adding a **small** amount of noise to price such that

$$p = \frac{A}{B} - \frac{D}{B} \times \text{average price forecast} + \textit{noise}.$$

The *noise* term is chosen at random in each period and is not predictable. Its value is not given to any participant in the market. The size of each realisation is **small**. The average value of the *noise* over the course of the experiment is zero and each realisation of it is independent from any other realization. In other words, the *noise* term may take a positive or a negative value in any given period, but overall, the size and number of positive and negative realisations will be approximately equal and cancel each other out over time.

Paper instructions:

Widget Game Instruction Summary:

- Your job is to forecast the price of a widget next period
- Demand for widgets is determined by the market price
  - $q = A - B\,p$
- The total supply of widgets to the market is determined by the **average** of all price forecasts submitted to the market
  - $q = D$ *x* average price forecast
- Combining supply and demand, we have the **key formula** that determines price in the market
  - $P = A/B - D/B$ *x* average expected price + *noise*
    - Recall that *noise* is small and **on average equal to zero**
- **An Example:** $A = 120$, $B = 2$, $D = 1$, and $noise = 0$, what is price if the average price forecast is 42?
  - $p = 60 - ½$ *x* average price forecast
  - $P = 60 - ½ \times 42 = 60 - 21 = 39$
- You are paid based on **accuracy of your forecast** according to the following formula
  - Payment = 0.50 – 0.03 $(p - your\ price\ forecast)^2$
    - A perfect forecast in a round earns 50 cents
    - A very poor forecast results in 0.00
- **KEY POINT: The market has negative feedback. Therefore, if the average price forecast is high, the market price will be low. And, if the average price forecast is low, then the market price will be high.**
- **Your Notes:**
  - –
  - –
  - –
  - –
  - –
  - –

Widget Game Rules

- You may withdraw from the experiment at any time for any reason
- You may take notes on this paper or the scratch paper provided
- Feel free to do any calculations you wish on the scratch paper provided
- **Do not exit the web browser**
- Do not open new tabs in the web browser
- **Please turn your phone off during the experiment**
- Do not speak with the people around you

Widget Game Instruction Summary:

- Your job is to forecast the price of a widget next period
- Demand for widgets is determined by the market price
  - *q = A + B p*
- The total supply of widgets to the market is determined by the **average** of all price forecasts submitted to the market
  - *q* = C + *D x* average price forecast
- Combining supply and demand, we have the **key formula** that determines price in the market
  - *P = (C – A)/B + D/B x* average expected price + *noise*
    - Recall that *noise* is small and **on average equal to zero**
- **An Example:** *A = 0, B =2, C=60, D = 1,* and *noise = 0*, what is price if the average price forecast is 42?
  - *p = 30 + ½ x* average price forecast
  - *P = 30 + ½ x 42 = 30 + 21 = 51*
- You are paid based on **accuracy of your forecast** according to the following formula
  - Payment = 0.50 – 0.03 $(p - your\ price\ forecast)^2$
    - A perfect forecast in a round earns 50 cents
    - A very poor forecast results in 0.00
- **KEY POINT: The market has positive feedback. Therefore, if the average price forecast is high, the market price will be high. And, if the average price forecast is low, then the market price will be low.**
- **Your Notes:**
  - –
  - –
  - –
  - –
  - –
  - –

Widget Game Rules

- You may withdraw from the experiment at any time for any reason
- You may take notes on this paper or the scratch paper provided
- Feel free to do any calculations you wish on the scratch paper provided
- **Do not exit the web browser**
- Do not open new tabs in the web browser
- **Please turn your phone off during the experiment**
- Do not speak with the people around you