Online Appendix for A Single-Judge Solution to Beauty Contests

By Zhen Huo and Marcelo Pedroni*

This appendix contains all proofs not included in the main text. To facilitate reading we include Appendices A and B which can also be found in the main text with the exception of the proof of Proposition 1.

Appendix A: Proof of Theorem 1

The proof is presented as a series of lemmas and propositions. In Section A1 we show that the forecasting problem set up in Section II is equivalent to a limit of a truncated version of it. Section A2 sets up and solves a static forecasting problem equivalent to the truncated version. Section A3 presents and solves the associated fixed point problem that gives the equilibrium of a beauty-contest problem; in it, we also establish the existence, uniqueness, and linearity of the equilibrium. Section A4 describes the α-modified signal process and Section A5 proves the equivalence between the policy function of the solution to the static version of the forecasting problem with the α-modified signal process and the solution to the fixed point problem.

A1. Limit of Truncated Forecasting Problem

Fix t. Section II sets up the problem of forecasting θ_t given $\mathbf{x}_i^t \equiv \{\mathbf{x}_{it}, \mathbf{x}_{it-1}, \ldots\}$. For ease of notation, we define

$$\vartheta \equiv \theta_t = \sum_{k=0}^{\infty} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k}, \quad \text{and} \quad \boldsymbol{x} \equiv \mathbf{x}_i^t.$$

Notice that each element of ϑ and $\boldsymbol{x}$ can be represented as an MA(∞) process and that there is an infinite history of signals.

Consider a truncated version of this problem.[1] Let ϑ_q be the MA(q) truncation of ϑ, that is,

$$\vartheta_q = \sum_{k=0}^{q} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k}.$$

Let $\boldsymbol{x}_p^{(N)} \equiv \{\mathbf{x}_{p,it}, \ldots, \mathbf{x}_{p,it-N}\}$ where $\mathbf{x}_{p,it-k}$ is the MA(p) truncation of $\mathbf{x}_{it-k}$. The next proposition shows that the limit as q, p, and N go to infinity of the forecast of ϑ_q given $\boldsymbol{x}_p^{(N)}$ is equivalent to the forecast of ϑ given $\boldsymbol{x}$. Throughout, the concept of convergence between random variables is mean square. For example, we can say that $\lim_{q\to\infty} \vartheta_q = \vartheta$, since

$$\lim_{q\to\infty} \mathbb{E}\left[(\vartheta - \vartheta_q)^2\right] = \lim_{q\to\infty} \mathbb{E}\left[\left(\vartheta - \sum_{k=0}^{q} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k}\right)^2\right] = \lim_{q\to\infty} \sum_{k=q+1}^{\infty} \boldsymbol{\phi}_k \mathbb{E}[\boldsymbol{\eta}_{t-k}^2] \boldsymbol{\phi}_k' = 0,$$

where the last equality is due to the assumption that $\boldsymbol{\phi}(L)$ is square summable and that $\mathbb{E}[\boldsymbol{\eta}_{t-k}^2]$ is finite.

PROPOSITION 1: $\mathbb{E}[\vartheta \mid \boldsymbol{x}] = \lim_{p,q,N\to\infty} \mathbb{E}\left[\vartheta_q \mid \boldsymbol{x}_p^{(N)}\right]$.

PROOF:

* Huo: Yale University, 28 Hillhouse Ave, New Haven, CT 06510, zhen.huo@yale.edu. Pedroni: University of Amsterdam, Roetersstraat 11, Amsterdam, 1018WB, m.pedroni@uva.nl.

[1] Note that in this truncation, we do not assume shocks become public after a certain number of periods, differently from the common assumption made in the literature (e.g. Townsend (1983)).

The strategy is to establish the following equalities

$$\mathbb{E}\left[\vartheta \mid \boldsymbol{x}\right] \overset{[3]}{=} \lim_{p\to\infty} \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p\right] \overset{[2]}{=} \lim_{p\to\infty} \lim_{N\to\infty} \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N)}\right] \overset{[1]}{=} \lim_{p\to\infty} \lim_{N\to\infty} \lim_{q\to\infty} \mathbb{E}\left[\vartheta_q \mid \boldsymbol{x}_p^{(N)}\right].$$

We start from the last and move to the first.

[1]: To show that

$$\mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N)}\right] = \lim_{q\to\infty} \mathbb{E}\left[\vartheta_q \mid \boldsymbol{x}_p^{(N)}\right],$$

note that there exists K large enough such that, for any $k > K$,

$$\mathbb{E}\left[\boldsymbol{\eta}_{t-k} \mid \boldsymbol{x}_p^{(N)}\right] = 0.$$

It follows that

$$\mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N)}\right] = \mathbb{E}\left[\sum_{k=0}^{K} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k} + \sum_{k=K+1}^{\infty} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k} \mid \boldsymbol{x}_p^{(N)}\right] = \mathbb{E}\left[\sum_{k=0}^{K} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k} \mid \boldsymbol{x}_p^{(N)}\right] = \mathbb{E}\left[\vartheta_K \mid \boldsymbol{x}_p^{(N)}\right],$$

and

$$\lim_{q\to\infty} \mathbb{E}\left[\vartheta_q \mid \boldsymbol{x}_p^{(N)}\right] = \lim_{q\to\infty} \mathbb{E}\left[\sum_{k=0}^{K} \boldsymbol{\phi}_k \boldsymbol{\varepsilon}_{-k} + \sum_{k=K+1}^{q} \boldsymbol{\phi}_k \boldsymbol{\varepsilon}_{-k} \mid \boldsymbol{x}_p^{(N)}\right] = \lim_{q\to\infty} \mathbb{E}\left[\vartheta_K \mid \boldsymbol{x}_p^{(N)}\right] = \mathbb{E}\left[\vartheta_K \mid \boldsymbol{x}_p^{(N)}\right].$$

[2]: Next, to show that

$$\mathbb{E}[\vartheta \mid \boldsymbol{x}_p] = \lim_{N\to\infty} \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N)}\right],$$

we simply need to establish that the limit on the right hand side exists. First notice that forecast errors are decreasing in the number of signals and that the stationarity of ϑ guarantees that the mean squared error is well defined, which implies that

$$0 \leq \mathbb{E}\left[\left(\vartheta - \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N+1)}\right]\right)^2\right] \leq \mathbb{E}\left[\left(\vartheta - \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N)}\right]\right)^2\right].$$

Therefore, there exists σ^2 such that

$$\lim_{N\to\infty} \mathbb{E}\left[\left(\vartheta - \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N)}\right]\right)^2\right] = \sigma^2.$$

Moreover, for any N_1, N_2,

$$\mathbb{E}\left[\left(\vartheta - \frac{\mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_1)}\right] + \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_2)}\right]}{2}\right)^2\right] \geq \sigma^2.$$

It follows that

$$\mathbb{E}\left[\left(\mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_1)}\right] - \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_2)}\right]\right)^2\right]$$
$$= 2\mathbb{E}\left[\left(\vartheta - \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_1)}\right]\right)^2\right] + 2\mathbb{E}\left[\left(\vartheta - \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_2)}\right]\right)^2\right] - 4\mathbb{E}\left[\left(\vartheta - \frac{\mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_1)}\right] + \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_2)}\right]}{2}\right)^2\right]$$

$$\leq 2\mathbb{E}\left[\left(\vartheta - \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_1)}\right]\right)^2\right] + 2\mathbb{E}\left[\left(\vartheta - \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N_2)}\right]\right)^2\right] - 4\sigma^2.$$

In the limit, the right-hand side converges to zero, and therefore

$$\lim_{N\to\infty} \mathbb{E}\left[\vartheta \mid \boldsymbol{x}_p^{(N)}\right]$$

is indeed well defined.

[3]: Recall that from equation (11),

$$\mathbf{x}_{it} = \mathbf{M}_t(L)\boldsymbol{\varepsilon}_{it}.$$

Let $\mathbf{M}_{p,t}$ be the MA (p) truncation of $\mathbf{M}_t(L)$, so that

$$\mathbf{x}_{p,it} = \mathbf{M}_{p,t}(L)\boldsymbol{\varepsilon}_{it}.$$

To obtain a formula for the forecast conditional on infinite signals, we use the Wiener-Hopf prediction formula (Whittle, 1963), which leads to[2]

$$\begin{aligned}
\mathbb{E}[\vartheta \mid \boldsymbol{x}] &= [\boldsymbol{\phi}(L)\mathbf{M}_t'(L^{-1})\mathbf{B}'(L^{-1})^{-1}]_+\mathbf{B}(L)^{-1}\mathbf{x}_{it} \equiv \mathbf{D}(L)\boldsymbol{\varepsilon}_{it}, \\
\mathbb{E}[\vartheta \mid \boldsymbol{x}_p] &= [\boldsymbol{\phi}(L)\mathbf{M}_{p,t}'(L^{-1})\mathbf{B}_p'(L^{-1})^{-1}]_+\mathbf{B}_p(L)^{-1}\mathbf{x}_{p,it} \equiv \mathbf{D}_p(L)\boldsymbol{\varepsilon}_{it},
\end{aligned} \tag{A1}$$

where $\mathbf{B}(z)$ and $\mathbf{B}_p(z)$ are the corresponding fundamental representations of $\mathbf{M}_t(z)$ and $\mathbf{M}_{p,t}(z)$, respectively. The mean-squared difference between these forecasts is

$$\mathbb{E}\left[(\mathbb{E}[\vartheta \mid \boldsymbol{x}] - \mathbb{E}[\vartheta \mid \boldsymbol{x}_p])^2\right] = \sum_{k=0}^{\infty}(\mathbf{D}_{t-k} - \mathbf{D}_{p,t-k})\boldsymbol{\Sigma}^2(\mathbf{D}_{t-k} - \mathbf{D}_{p,t-k})'.$$

It is, therefore, sufficient to show that $\lim_{p\to\infty}\mathbf{D}_p(z) = \mathbf{D}(z)$, since it then follows that

$$\lim_{p\to\infty} \mathbb{E}\left[(\mathbb{E}[\vartheta \mid \boldsymbol{x}] - \mathbb{E}[\vartheta \mid \boldsymbol{x}_p])^2\right] = 0,$$

which establishes the result. To see that this is the case, notice that, given a signal process $\mathbf{M}_{p,t}(z)$, its corresponding fundamental representation of $\mathbf{B}_p(z)$ is uniquely determined when the covariance matrix of the fundamental innovation is normalized, and satisfies

$$\mathbf{B}_p(z)\mathbf{B}_p'(z^{-1}) = \mathbf{M}_{p,t}(z)\mathbf{M}_{p,t}'(z^{-1}).$$

As a result, since by construction we have that

$$\mathbf{M}_t(z) = \lim_{p\to\infty}\sum_{k=0}^{p}\mathbf{M}_{k,t}z^k = \lim_{p\to\infty}\mathbf{M}_{p,t}(z),$$

it follows that

$$\mathbf{B}(z) = \lim_{p\to\infty}\mathbf{B}_p(z).$$

Finally, by continuity of the annihilation operator, $[\,\cdot\,]_+$, in equation (A1), we obtain

$$\lim_{p\to\infty}\mathbf{D}_p(z) = \mathbf{D}(z).$$

[2] Notice that, if we use the standard prediction formula (used in Section A2 below), the forecast conditional on the infinite history of signals $\boldsymbol{x}$ would involve the product of infinitely dimensional matrices.

A2. Static Forecasting Problem

We now consider the truncated problem of forecasting ϑ_q conditional on $\boldsymbol{x}_p^{(N)}$, which can be viewed as a static problem. Again, to ease notation, define

$$\boldsymbol{\eta} \equiv \begin{bmatrix} \boldsymbol{\eta}_t \\ \vdots \\ \boldsymbol{\eta}_{t-T} \end{bmatrix}, \quad \boldsymbol{\nu}_i \equiv \begin{bmatrix} \boldsymbol{\nu}_{it} \\ \vdots \\ \boldsymbol{\nu}_{it-T} \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\varepsilon}_i \equiv \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\nu}_i \end{bmatrix},$$

where $T \equiv \max\{q, p+N\}$. Notice that, there exists a vector $\boldsymbol{a}$ with length $U \equiv u(T+1)$, and a matrix $\mathbf{B}$ with dimensions $r(N+1) \times M$, where $M \equiv m(T+1)$, such that the forecasting problem at time t becomes that of forecasting

$$\theta \equiv \vartheta_q = \begin{bmatrix} \mathbf{a}' & \mathbf{0}' \end{bmatrix} \boldsymbol{\varepsilon}_i = \mathbf{a}'\boldsymbol{\eta}, \quad \text{given} \quad \mathbf{x}_i \equiv \begin{bmatrix} \mathbf{x}_{p,it} \\ \vdots \\ \mathbf{x}_{p,it-N} \end{bmatrix} = \mathbf{B}\boldsymbol{\varepsilon}_i = \mathbf{B} \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\nu}_i \end{bmatrix}.$$

Let $\mathbf{\Omega}^2$ denote the covariance matrix of $\boldsymbol{\varepsilon}_i$, let $\mathbf{A} \equiv \begin{bmatrix} \mathbf{a}' & \mathbf{0}' \end{bmatrix}'$, and let $\mathbf{\Lambda}$ be the $M \times M$ matrix given by

$$\mathbf{\Lambda} \equiv \begin{bmatrix} \mathbf{I}_U & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

It follows that

$$\mathbb{E}[\theta \mid \mathbf{x}_i] = \mathbf{A}'\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}\mathbf{B}' \left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{x}_i.$$

It is convenient for what follows to write the forecast in this way. To obtain this formula we use, in particular, the fact that $\mathbf{A}' = \mathbf{A}'\mathbf{\Lambda}$ and that $\mathbf{\Lambda}\mathbf{\Omega} = \mathbf{\Omega}\mathbf{\Lambda}$.

A3. Fixed Point Problem

Suppose we also want to forecast y. We do not know the stochastic process for y, so let $\mathbf{h}$ denote the agent's equilibrium policy function, i.e. $y_i = \mathbf{h}'\mathbf{x}_i$, then

$$y = \int \mathbf{h}'\mathbf{x}_i = \mathbf{h}'\mathbf{B}\mathbf{\Lambda}\boldsymbol{\varepsilon}_i.$$

Then, the forecast of y is given by

$$\mathbb{E}[y \mid \mathbf{x}_i] = \mathbf{h}'\mathbf{B}\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}\mathbf{B}' \left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{x}_i.$$

In equilibrium, we have that

$$y_i = (1-\alpha)\mathbb{E}[\theta \mid \mathbf{x}_i] + \alpha\mathbb{E}[y \mid \mathbf{x}_i],$$

and, therefore

$$\mathbf{h}'\mathbf{x}_i = \left[(1-\alpha)\mathbf{A}'\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}\mathbf{B}' \left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1} + \alpha\mathbf{h}'\mathbf{B}\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}\mathbf{B}' \left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1}\right] \mathbf{x}_i.$$

It follows from the fact that equation above holds for any $\mathbf{x}_i$ that

$$\mathbf{h} = \mathbf{C}^{-1}\mathbf{d}, \tag{A2}$$

where

$$\mathbf{C} \equiv \mathbf{I} - \alpha\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}', \tag{A3}$$

$$\mathbf{d} \equiv (1-\alpha)\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{A}. \tag{A4}$$

LEMMA A.1: $\mathbf{C}$ *is invertible.*

PROOF:
We start by showing that $\mathbf{Y} \equiv \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}'$ has real eigenvalues in $[0,1]$. First notice that $\mathbf{Y}$ has real, non-negative eigenvalues since it is similar to

$$\begin{aligned}(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{1/2}\mathbf{Y}(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1/2} &= (\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{1/2}(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1}(\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}')(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1/2}\\ &= (\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1/2}(\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}')(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1/2}\end{aligned}$$

which is positive semidefinite. On the other hand,

$$\mathbf{I} - \mathbf{Y} = \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}(\mathbf{I}-\boldsymbol{\Lambda})\boldsymbol{\Omega}\mathbf{B}',$$

which, analogously to $\mathbf{Y}$, is also similar to a positive semidefinite matrix. If λ is an eigenvalue of $\mathbf{Y}$, then $1-\lambda$ is an eigenvalue of $\mathbf{I}-\mathbf{Y}$. Therefore, the fact that the eigenvalues of $\mathbf{I}-\mathbf{Y}$ are positive implies that the eigenvalues of $\mathbf{Y}$ must be less than or equal to 1, as desired. It follows that $\mathbf{S} \equiv \sum_{j=0}^{\infty}(\alpha\mathbf{Y})^j$ converges, and $\mathbf{S}\left(\mathbf{I}-\alpha\mathbf{Y}\right) = \mathbf{I}$.

In particular, it follows from this lemma that there exists a unique equilibrium to the beauty-contest model. It also follows that the equilibrium actions of an agent are linear functions of their signals.[3]

A4. α-Modified Signal Process and Prediction Formula

Define $\boldsymbol{\Gamma}$ to be the $M \times M$ matrix given by

$$\boldsymbol{\Gamma} \equiv \begin{bmatrix} \mathbf{I}_U & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{1-\alpha}}\mathbf{I}_{M-U} \end{bmatrix},$$

and suppose that the signals observed by agent i are a α-modified version of $\mathbf{x}_i$ given by

$$\widetilde{\mathbf{x}}_i = \mathbf{B}\widetilde{\boldsymbol{\varepsilon}}_i, \quad \text{with} \quad \widetilde{\boldsymbol{\varepsilon}}_i \equiv \boldsymbol{\Gamma}\boldsymbol{\varepsilon}_i.$$

It follows that

$$\widetilde{\mathbb{E}}[\theta \mid \widetilde{\mathbf{x}}_i] = \mathbf{A}'\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}'\left(\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Gamma}^2\boldsymbol{\Omega}\mathbf{B}'\right)^{-1}\widetilde{\mathbf{x}}_i. \tag{A5}$$

A5. Equivalence Result

Proposition 2 establishes that the right hand side of equation terms in equation (A2) is to the right hand side of equation (A5), which completes the proof of Theorem 1.

PROPOSITION 2: *Using the definitions above, it follows that*

$$\mathbf{C}^{-1}\mathbf{d} = \left(\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Gamma}^2\boldsymbol{\Omega}\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{A}.$$

[3] In the arguments made above we have implicitly used the well known result that the optimal forecast for Gaussian processes is linear, see Hamilton (1994) Section 4.6 for a formal proof.

PROOF:

From the definition of $\mathbf{C}$, equation (A3), we obtain

$$\begin{aligned}\mathbf{C} &= \mathbf{I} - \alpha \left(\mathbf{B\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B\Omega\Lambda\Omega B}' \\ &= \left(\mathbf{B\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B\Omega}(\mathbf{I} - \alpha\mathbf{\Lambda})\mathbf{\Omega B}'.\end{aligned}$$

Thus, since

$$\mathbf{I} - \alpha\mathbf{\Lambda} = (1-\alpha)\mathbf{\Gamma}^2,$$

it follows that

$$\mathbf{C} = (1-\alpha)\left(\mathbf{B\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B\Omega\Gamma}^2\mathbf{\Omega B}'.$$

Finally, using Lemma A.1 and equation (A4),

$$\begin{aligned}\mathbf{C}^{-1}\mathbf{d} &= \left[(1-\alpha)\left(\mathbf{B\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B\Omega\Gamma}^2\mathbf{\Omega B}'\right]^{-1} (1-\alpha)\left(\mathbf{B\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B\Omega\Lambda\Omega A} \\ &= (\mathbf{B\Omega\Gamma}^2\mathbf{\Omega B}')^{-1}\left(\mathbf{B\Omega}^2\mathbf{B}'\right)\left(\mathbf{B\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B\Omega\Lambda\Omega A} \\ &= (\mathbf{B\Omega\Gamma}^2\mathbf{\Omega B}')^{-1}\mathbf{B\Omega\Lambda\Omega A}.\end{aligned}$$

REMARK

It is crucial for the proof to work that the matrix $\mathbf{I}-\alpha\mathbf{\Lambda}$ is real, symmetric and positive semidefinite so that it can be interpreted as a covariance matrix. When there are forward-looking or backward-looking strategic complementarities this is in general not the case. This does not mean that a transformation to the information structure that yields an equivalence result cannot exist with dynamic complementarities, but simply that this particular proof strategy is not suitable in those cases. See, for instance, Proposition 8.

Appendix B: Proofs for Generalized Best Response and Multiple Actions

This appendix contains the proofs for the extensions to generalized best responses and multiple actions from Section III.C.

B1. Proof of Proposition 3

PROOF:
Iterating on the best response function in equation (17) we obtain

$$y_{it} = \gamma(\varphi_t + \xi_{it}) + \alpha\gamma \sum_{k=0}^{\infty} \alpha^k \, \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k\left[\varphi_t\right]\right].$$

By Corollary 2, the infinite sum of higher-order expectations can be rewritten as a first-order expectation

$$y_{it} = \gamma(\varphi_t + \xi_{it}) + \frac{\alpha\gamma}{1-\alpha}\mathbf{h}_t'(L)\mathbf{x}_{it}.$$

B2. Proof of Proposition 4

PROOF:
Iterating on equation (18) leads to

$$\mathbf{y}_{it} = \sum_{k=0}^{\infty} \mathbf{A}^k \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k\left[\boldsymbol{\theta}_t\right]\right].$$

Then, notice that for any $k \in \{0, 1, 2, \ldots\}$,

$$\mathbf{A}^k = \mathbf{Q}\,\mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)^k \mathbf{Q}^{-1} = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'\mathbf{Q}^{-1}\alpha_j^k.$$

Therefore, $\mathbf{y}_{it}$ can be written as

$$\mathbf{y}_{it} = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'\mathbf{Q}^{-1} \sum_{k=0}^{\infty} \alpha_j^k \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k\left[\boldsymbol{\theta}_t\right]\right].$$

From Corollary 2 we have that each row of $\sum_{k=0}^{\infty} \alpha_j^k \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k\left[\boldsymbol{\theta}_t\right]\right]$ is equal to the corresponding row of $(1-\alpha_j)^{-1}\mathbf{g}_{jt}'(L)\mathbf{x}_{it}$ and it follows that

$$\mathbf{y}_{it} = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'(1-\alpha_j)^{-1}\mathbf{Q}^{-1}\mathbf{g}_{jt}'(L)\mathbf{x}_{it} = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'\left(\mathbf{I} - \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)\right)^{-1}\mathbf{Q}^{-1}\mathbf{g}_{jt}'(L)\mathbf{x}_{it},$$

and the fact that $\left(\mathbf{I} - \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)\right)^{-1}\mathbf{Q}^{-1} = \mathbf{Q}^{-1}\left(\mathbf{I} - \mathbf{A}\right)^{-1}$ concludes the proof.

Appendix C: Proof of Proposition 5

PROOF:

By Proposition 2, the individual action can be written as

$$y_{it} = \mathbf{G}\mathbf{z}_{it},$$

where

$$\mathbf{z}_{it} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{F}\mathbf{z}_{it-1} + \mathbf{K}\mathbf{x}_{it} = (\mathbf{I} - (\mathbf{F} - \mathbf{K}\mathbf{H}\mathbf{F})L)^{-1}\mathbf{K}\mathbf{x}_{it},$$

and $\mathbf{K}$ is the steady state Kalman gain matrix with α-modified signals. Note that by Cramer's rule,

$$(\mathbf{I} - (\mathbf{F} - \mathbf{K}\mathbf{H}\mathbf{F})L)^{-1} = \frac{\mathrm{adj}(\mathbf{I} - (\mathbf{F} - \mathbf{K}\mathbf{H}\mathbf{F})L)}{\det(\mathbf{I} - (\mathbf{F} - \mathbf{K}\mathbf{H}\mathbf{F})L)} = \frac{\mathrm{adj}(\mathbf{I} - (\mathbf{F} - \mathbf{K}\mathbf{H}\mathbf{F})L)}{\Pi_{k=1}^{\ell}(1 - \lambda_k L)},$$

where $\{\lambda_k\}_{k=1}^{\ell}$ are non-zero eigenvalues of $(\mathbf{F} - \mathbf{K}\mathbf{H}\mathbf{F})$. Let r denote the dimension of $\mathbf{F}$, the last equality follows from the fact that

$$\det\left(\mathbf{I} - (\mathbf{F} - \mathbf{K}\mathbf{H}\mathbf{F})L\right) = L^r \det\Big(\mathbf{I}L^{-1} - (\mathbf{F} - \mathbf{K}\mathbf{H}\mathbf{F})\Big) = L^r L^{-(r-\ell)} \prod_{k=1}^{\ell}(L^{-1} - \lambda_k) = \prod_{k=1}^{\ell}(1 - \lambda_k L).$$

The aggregate outcome y_t then follows

$$y_t = \int y_{it} = \mathbf{C}(L) \int \mathbf{x}_{it} = \mathbf{C}(L)\Big(\mathbf{A}\theta_t + \mathbf{B}(L)\mathbf{v}_t\Big).$$

Appendix D: Variance and Covariance of Average Forecast

The following proof makes use of the same limiting arguments used in Appendix A1, which allows us to focus on the truncated static problem. Let $\boldsymbol{\Omega} = \operatorname{diag}(\sigma_1, \ldots, \sigma_M)$, and $\mathcal{I} \equiv \{U+1, \ldots, M\}$ be the σ-indexes associated with the idiosyncratic shocks $\boldsymbol{\nu}_i$. Also, let $\mathbf{e}_j$ be the j-th column of the $M \times M$ identity matrix and let $\overline{\mathbb{E}}[\theta] \equiv \int \mathbb{E}[\theta \mid \mathbf{x}_i]$.

Increasing the variance of any element of $\boldsymbol{\nu}_i$ does not affect θ, but makes the signals $\mathbf{x}_i$ noisier which, in turn, makes the forecast less accurate. This implies that the forecast reacts less to signals and, as a result, it is less volatile. It also implies that it is less correlated with the actual θ. This motivates the following lemmas.

LEMMA D.1: *The variance* $\operatorname{Var}(\overline{\mathbb{E}}[\theta])$ *is decreasing in how noisy the signals are,*

$$\frac{\partial \operatorname{Var}(\overline{\mathbb{E}}[\theta])}{\partial \sigma_j^2} \leq 0, \quad \text{for } j \in \mathcal{I}.$$

PROOF:

First notice that

$$\operatorname{Var}(\overline{\mathbb{E}}[\theta]) = \mathbf{A}'\boldsymbol{\Omega}^2\mathbf{B}'\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}'\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}^2\mathbf{A}.$$

For any $j \in \mathcal{I}$, $\mathbf{A}'\boldsymbol{\Omega}^2$ and $\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}$ do not depend on σ_j^2, and therefore, taking derivatives yields

$$\begin{aligned}\frac{\partial \operatorname{Var}(\overline{\mathbb{E}}[\theta])}{\partial \sigma_j^2} = &- \mathbf{A}'\boldsymbol{\Omega}^2\mathbf{B}'\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\Bigg[\left(\mathbf{B}\mathbf{e}_j\mathbf{e}_j'\mathbf{B}'\right)\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}' + \\ &+ \mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}'\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\left(\mathbf{B}\mathbf{e}_j\mathbf{e}_j'\mathbf{B}'\right)\Bigg]\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}^2\mathbf{A}.\end{aligned}$$

The matrix in the inner bracket is symmetric, so let $\mathbf{L}\mathbf{L}'$ denote its Cholesky decomposition. Then, letting $\mathbf{z} \equiv \mathbf{A}'\boldsymbol{\Omega}^2\mathbf{B}'\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{L}$, the right hand side is equal to $-\sigma_j^2\mathbf{z}\mathbf{z}'$, which is less than or equal to 0.

LEMMA D.2: *The covariance* $\operatorname{Cov}(\theta, \overline{\mathbb{E}}[\theta])$ *is decreasing in how noisy the signals are,*

$$\frac{\partial \operatorname{Cov}(\theta, \overline{\mathbb{E}}[\theta])}{\partial \sigma_j^2} \leq 0, \quad \text{for } j \in \mathcal{I}.$$

PROOF:

First notice that

$$\operatorname{Cov}(\theta, \overline{\mathbb{E}}[\theta]) = \mathbf{A}'\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}'\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}^2\mathbf{A}.$$

For any $j \in \mathcal{I}$, $\mathbf{A}'\boldsymbol{\Omega}^2$ does not depend on σ_j^2, and $\mathbf{A}'\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega} = \mathbf{A}'\boldsymbol{\Omega}^2$, therefore

$$\frac{\partial \operatorname{Cov}(\theta, \overline{\mathbb{E}}[\theta])}{\partial \sigma_j^2} = -\mathbf{A}'\boldsymbol{\Omega}^2\mathbf{B}'\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\left(\mathbf{B}\mathbf{e}_j\mathbf{e}_j'\mathbf{B}'\right)\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}^2\mathbf{A}.$$

Finally, letting $\mathbf{z} \equiv \mathbf{A}'\boldsymbol{\Omega}^2\mathbf{B}'\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\mathbf{e}_j$, the right hand side of this equation can be written as $-\mathbf{z}\mathbf{z}'$ which is less than or equal to 0.

For the following proofs we denote $\widetilde{\mathbb{E}}[\theta] \equiv \int \mathbb{E}[\theta \mid \widetilde{\mathbf{x}}_i]$, where $\widetilde{\mathbf{x}}_i$ is the modified signal defined in Section A4.

D1. Proof of Proposition 6 for the Common-Value Best Response

PROOF:

In the static version of the problem, under Assumptions 1-3 and with the best response given by

$$y_i = (1-\alpha)\mathbb{E}_i[\theta] + \alpha\mathbb{E}_i[y],$$

it follows from Corollary 1 that

$$y = \widetilde{\mathbb{E}}[\theta],$$

which is equivalent to $\overline{\mathbb{E}}[\theta]$ with the variance of the idiosyncratic shocks discounted by α, that is with $\widetilde{\sigma}_j^2 = \sigma_j^2/(1-\alpha)$ for all $j \in \mathcal{I}$. Hence, an increase in α is equivalent to an increase in the variance of all the idiosyncratic shocks. Thus, Part 1 in Proposition 6 for the CVBR follows from Lemmas D.1 and D.2. Part 2 follows directly from Lemma D.1 and the fact that $y = \overline{\mathbb{E}}[\theta]$ when $\alpha = 0$. Finally, it follows from Proposition 1 that these results generalize to the setting in Proposition 6 by a continuity argument.

D2. Proof of Proposition 6 for the Independent-Value Best Response

PROOF:

With the best response given by

$$y_i = (1-\alpha)\theta_i + \alpha\mathbb{E}_i[y],$$

it follows from Proposition 3 that

$$y = (1-\alpha)\theta + \alpha\widetilde{\mathbb{E}}[\theta].$$

Notice that

$$\mathrm{Var}(y) = (1-\alpha)^2\mathrm{Var}(\theta) + \alpha^2\mathrm{Var}(\widetilde{\mathbb{E}}[\theta]) + 2\alpha(1-\alpha)\mathrm{Cov}(\theta, \widetilde{\mathbb{E}}[\theta]).$$

Then, using the law of total variance and Lemmas D.1 and D.2, it follows that, if $\alpha > 0$,

$$\begin{aligned}\frac{\partial \mathrm{Var}(y)}{\partial \alpha} =& -2(1-\alpha)\mathrm{Var}(\theta) + 2\alpha\mathrm{Var}(\widetilde{\mathbb{E}}[\theta]) + 2(1-2\alpha)\mathrm{Cov}(\theta, \widetilde{\mathbb{E}}[\theta])\\ &+ \alpha^2\frac{\partial \mathrm{Var}(\widetilde{\mathbb{E}}[\theta])}{\partial \alpha} + 2\alpha(1-\alpha)\frac{\partial \mathrm{Cov}(\theta, \widetilde{\mathbb{E}}[\theta])}{\partial \alpha}\\ \leq& -2(1-\alpha)\mathrm{Var}(\theta) + 2\alpha\mathrm{Var}(\theta) + 2(1-2\alpha)\mathrm{Var}(\theta) = 0.\end{aligned}$$

Similarly,

$$\mathrm{Cov}(y, \theta) = (1-\alpha)\mathrm{Var}(\theta) + \alpha\mathrm{Cov}(\theta, \widetilde{\mathbb{E}}[\theta]),$$

and if $\alpha > 0$,

$$\begin{aligned}\frac{\partial \mathrm{Cov}(y, \theta)}{\partial \alpha} &= -\mathrm{Var}(\theta) + \mathrm{Cov}(\theta, \widetilde{\mathbb{E}}[\theta]) + \alpha\frac{\partial \mathrm{Cov}(\theta, \widetilde{\mathbb{E}}[\theta])}{\partial \alpha}\\ &\leq -\mathrm{Var}(\theta) + \mathrm{Var}(\theta) = 0.\end{aligned}$$

This establishes Part 1 for IVBR. Part 3 for $\alpha > 0$ follows from the fact that $y_t = \theta_t$ when $\alpha = 0$ and the result in Part 1. Finally, we can also write the aggregate action y as

$$y = \theta + \alpha(\widetilde{\mathbb{E}}[\theta] - \theta).$$

and it follows that

$$\mathrm{Var}(y) = \mathrm{Var}(\theta) + \alpha^2\mathrm{Var}(\widetilde{\mathbb{E}}[\theta] - \theta) + 2\alpha\,\mathrm{Cov}(\theta,\, \widetilde{\mathbb{E}}[\theta] - \theta).$$

To show that, for $\alpha < 0$, $\mathrm{Var}(y) \geq \mathrm{Var}(\theta)$, it is sufficient to show that $\mathrm{Cov}(\theta, \widetilde{\mathbb{E}}[\theta] - \theta) \leq 0$. Note that

$$\mathrm{Cov}(\theta, \widetilde{\mathbb{E}}[\theta] - \theta) = \mathrm{Cov}(\theta, \widetilde{\mathbb{E}}[\theta]) - \mathrm{Var}(\theta) \leq 0.$$

Proposition 1 implies that these results generalize to the setting in Proposition 6 by continuity.

Appendix E: Example Economies

In this appendix, we describe two different economic environments that have equilibria that can be summarized by a system of equations with a best response equation and an aggregation equation of the kind we work with in the paper.

E1. Monetary Model with Dispersed Information

In this section, we describe a simple monetary model with information frictions; it is a version of the model in Woodford (2002). A representative household has period utility given by

$$U(C_t, N_t) = \frac{C_t^{1-\sigma}}{1-\sigma} - \chi \frac{N_t^{1+\kappa}}{1+\kappa}$$

where N_t denotes their labor supply and C_t is their consumption of a composite good defined to be the CES aggregator of a continuum of differentiated goods,

$$C_t = \left(\int C_{it}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}.$$

The demand for good i and supply of labor in period t are given by

$$C_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\eta} C_t, \quad \text{and} \quad N_t = \left(\frac{W_t}{\chi P_t C_t^{\sigma}} \right)^{\frac{1}{\kappa}}, \qquad \text{with} \quad P_t \equiv \left(\int P_{it}^{1-\eta} \right)^{\frac{1}{1-\eta}},$$

and where W_t denotes the wage. There is continuum of firms, each producing one of the differentiated goods with the following production function,

$$C_{it} = AN_{it}^{\varepsilon}.$$

Firms have private information about the state of the world and, in period t, solve

$$\max_{P_{it}, C_{it}, N_{it}} \mathbb{E}_{it}[P_{it}C_{it} - W_t N_{it}], \qquad \text{subject to} \quad C_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\eta} C_t, \quad \text{and} \quad C_{it} = A_t N_{it}^{\varepsilon}.$$

It follows that

$$P_{it} = \frac{\eta}{\eta - 1} \frac{\mathbb{E}_{it} \left[W_t C_{it}^{\frac{1-\varepsilon}{\varepsilon}} \right]}{\varepsilon A^{\frac{1}{\varepsilon}}}.$$

Nominal GDP is determined exogenously by a monetary shock, Θ_t, so that

$$P_t C_t = \Theta_t.$$

Using this equation and the household's optimality conditions we obtain

$$P_{it} = \left(\frac{\chi\eta}{A^{\frac{1}{\varepsilon}} \varepsilon (\eta - 1)} \right)^{\frac{\varepsilon}{\eta + \varepsilon(1-\eta)}} \mathbb{E}_{it} \left[\Theta_t^{\frac{(\sigma-1)\varepsilon+1}{\eta+\varepsilon(1-\eta)}} P_t^{1 - \frac{(\sigma-1)\varepsilon+1}{\eta+\varepsilon(1-\eta)}} N_t^{\frac{\kappa\varepsilon}{\eta+\varepsilon(1-\eta)}} \right].$$

Let lower-case variables denote log-deviations from steady state. Integrating the production function we obtain as first order approximation (which is exact if all shocks are log-normal) that $c_t = \varepsilon n_t$, and it follows that

$$p_{it} = (1 - \alpha) \mathbb{E}_{it}[\theta_t] + \alpha \mathbb{E}_{it}[p_t], \qquad \text{with} \quad p_t = \int p_{it},$$

and the degree of strategic complementarity given by

$$\alpha \equiv \frac{(1-\varepsilon)(\eta-1)+\varepsilon(1-\sigma)-\kappa}{\eta+\varepsilon(1-\eta)}.$$

E2. Business Cycles Model

In this section, we describe a stylized real business cycle model with information frictions, a simplified version of the model in Angeletos and La'O (2010).

Environment

There is a continuum of islands indexed by i in the economy. In each island lives a representative agent who specializes in producing differentiated good i. Each period, agent i consumes C_{ijt} of the good produced on island j, and C_{it} is the CES aggregator of agent i's consumption of all goods,

$$C_{it} = \left(\int_j C_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}.$$

The production technology is

$$Y_{it} = \Theta_{it} N_{it}^{\varepsilon},$$

where η_{it} is the productivity level on island i in period t, and N_{it} denotes the labor input by agent i. The period utility of agent i is given by

$$U(C_{it}, N_{it}) = \frac{C_{it}^{1-\sigma}}{1-\sigma} - \frac{N_{it}^{1+\kappa}}{1+\kappa}.$$

Each period has two stages: In the first stage, agents decide how much to produce, that is, choose N_{it} which determines Y_{it}, conditional on their information about the island-specific productivity and aggregate output. In the second stage, taking prices and Y_{it} as given, they choose how much to consume of each good, i.e., $(C_{ijt})_j$, subject to the budget constraint

$$\int_j P_{jt} C_{ijt} = P_{it} Y_{it},$$

where P_{jt} is the price of the good produced on island j. Trading, in this second stage, occurs in a centralized market. In the first stage, agents have to decide how much to produce before the goods market opens, and therefore they have to forecast aggregate output to infer the price of their own goods. We allow agents to have different information sets which include their own productivity Θ_{it}, but otherwise, we remain agnostic about the information structure.

Equilibrium Characterization

Beginning with second stage, the optimal demand of the representative agent i for the good from island j whose price is P_{jt} is given by

$$C_{ijt} = C_{it}\left(\frac{P_{jt}}{P_t}\right)^{-\eta}, \qquad \text{where} \quad P_t \equiv \left(\int P_{jt}^{1-\eta}\right)^{\frac{1}{1-\eta}}.$$

Together with the budget constraint and market clearing condition $\int_i C_{ijt} = Y_{jt}$, it follows that

$$C_{it} = Y_t^{\frac{1}{\eta}} Y_{it}^{\frac{\eta-1}{\eta}} \qquad \text{where} \quad Y_t \equiv \left(\int Y_{jt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}},$$

that is, the consumption of island i is a weighted geometric mean of the aggregate output and the output produced in the island. Using this equation and the production function the first-stage problem becomes

$$\max_{Y_{it}} \mathbb{E}_{it}\left[\frac{1}{1-\sigma}\left(Y_t^{\frac{1}{\eta}} Y_{it}^{\frac{\eta-1}{\eta}}\right)^{1-\sigma} - \frac{1}{1+\kappa}\left(\frac{Y_{it}}{\Theta_{it}}\right)^{\frac{1+\kappa}{\varepsilon}}\right].$$

which implies

$$Y_{it} = \left(\varepsilon\frac{\eta-1}{\eta}\Theta_{it}^{\frac{1+\kappa}{\varepsilon}} \mathbb{E}_{it}\left[Y_t^{\frac{1-\sigma}{\eta}}\right]\right)^{\frac{\varepsilon\eta}{\eta(1+\kappa)+\varepsilon(1-\eta)(1-\sigma)}}.$$

Letting lower-case letters denote log-deviations from steady state, it follows that

$$y_{it} = \gamma\theta_{it} + \alpha\mathbb{E}_{it}[y_t], \qquad \text{with} \quad y_t = \int y_{it},$$

and the following definitions

$$\alpha \equiv \frac{\varepsilon\,(1-\sigma)}{\eta\,(1+\kappa)+\varepsilon\,(1-\eta)\,(1-\sigma)}, \qquad \text{and} \qquad \gamma \equiv \alpha\frac{\eta(1+\kappa)}{\varepsilon(1-\sigma)}.$$

This type of best response function is analyzed in Section III.C.

Appendix F: Proof of Theorem 2

The proof is presented as a series of lemmas and propositions. An analogous argument to the one made in Section A1 holds so that, if we prove the result for an arbitrary static information structure, the result follows. Section F1 sets up and solves the aforementioned arbitrary static forecasting problem. Section F2 presents and solves the associated fixed point problem that gives the equilibrium of a beauty-contest problem. Section F3 describes the modified signal process and Section F4 proves the equivalence between the policy function of the solution to the static version of the forecasting problem with the modified signal process and the solution to the fixed point problem.

F1. Setup with Static Information Structure

Best response function

The best response function is

$$\mathbf{y} = \mathbb{E}[\boldsymbol{\theta}] + \mathbb{E}[\mathbf{W}\mathbf{y}],$$

where $\mathbf{y}$ is a vector of individual actions

$$\mathbf{y} \equiv \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}',$$

$\boldsymbol{\theta}$ is a vector of exogenous variables

$$\boldsymbol{\theta} \equiv \begin{bmatrix} \theta_1 & \theta_2 & \cdots & \theta_n \end{bmatrix}',$$

and the expectation operator $\mathbb{E}$ is given by

$$\mathbb{E} \equiv \begin{bmatrix} \mathbb{E}_1 & \mathbb{E}_2 & \cdots & \mathbb{E}_n \end{bmatrix}'.$$

Agents can have heterogeneous information sets, and $\mathbb{E}_i$ may be different from $\mathbb{E}_j$. The matrix $\mathbf{W}$ represents the network structure.

Information structure

Let $\boldsymbol{\varepsilon}_i$ be the vector of, normally distributed, shocks the first u being common shocks, and the last r private. The covariance matrix of $\boldsymbol{\varepsilon}_i$ is the identity matrix. Suppose that the process for fundamentals is given by

$$\theta_i = \begin{bmatrix} \boldsymbol{\phi}_i' & \mathbf{0}_{1\times r} \end{bmatrix} \boldsymbol{\varepsilon}_i$$

and that agent i's signal is given by

$$\mathbf{x}_i = \mathbf{M}_i \boldsymbol{\varepsilon}_i$$

where

$$\mathbf{M}_i \equiv \mathbf{M}\boldsymbol{\Sigma}_i,$$

and

$$\boldsymbol{\Sigma}_i \equiv \operatorname{diag}\left(\tau_1^{-1/2}, \ldots, \tau_u^{-1/2}, \tau_{1i}^{-1/2}, \ldots, \tau_{ri}^{-1/2}\right),$$

with $u + r \equiv m$. Moreover, let

$$\boldsymbol{\Delta} \equiv \operatorname{diag}(\tau_1^{-1/2}, \ldots, \tau_u^{-1/2}, 0, \ldots, 0),$$

and define $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$ to be

$$\boldsymbol{\Lambda} \equiv \begin{bmatrix} \mathbf{I}_u \\ \mathbf{0}_{r\times u} \end{bmatrix}_{m\times u}, \quad \text{and} \quad \boldsymbol{\Gamma} \equiv \mathbf{I}_m - \boldsymbol{\Lambda}\boldsymbol{\Lambda}',$$

so that, in particular, we have

$$\mathbf{\Sigma}_i \mathbf{\Lambda} = \mathbf{\Delta}\mathbf{\Lambda}, \quad \text{and} \quad \mathbf{\Sigma}_i \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{\Sigma}_j = \mathbf{\Delta}^2.$$

Let $\mathbf{E}_i^n$ be the $n \times n$ matrix with zeros everywhere and 1 at the position (i,i), and define

$$\overline{\mathbf{M}} = \sum_{i=1}^{n} \mathbf{E}_i^n \otimes \mathbf{M}_i, \quad \text{and} \quad \mathbf{\Sigma} = \sum_{i=1}^{n} \mathbf{E}_i^n \otimes \mathbf{\Sigma}_i.$$

Finally, for each private shock, indexed by p, collect the associate variance for each agent i in the diagonal of the following matrix:

$$\mathbf{\Upsilon}_p \equiv \operatorname{diag}\left(\tau_{p1}^{-1}, \tau_{p2}^{-1}, \ldots, \tau_{pn}^{-1}\right).$$

Forecast

The forecast of the fundamental of agent j by agent i is given by

$$\mathbb{E}[\theta_j \mid \mathbf{x}_i] = \boldsymbol{\phi}_j' \mathbf{\Lambda}' \mathbf{M}_i' \left(\mathbf{M}_i \mathbf{M}_i'\right)^{-1} \mathbf{x}_i \equiv \mathbf{g}_{ji}' \mathbf{x}_i. \tag{F1}$$

F2. Fixed Point Problem

Let $\mathbf{h}_i$ be the equilibrium policy function for agent i, that is

$$y_i = \mathbf{h}_i' \mathbf{x}_i = \mathbf{h}_i' \mathbf{M}_i \boldsymbol{\varepsilon}_i.$$

Then, since the agents do not have information about others' private shocks, the forecast of y_j by agent i is given by

$$\mathbb{E}[y_j \mid \mathbf{x}_i] = \mathbf{h}_j' \mathbf{M}_j \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_i' \left(\mathbf{M}_i \mathbf{M}_i'\right)^{-1} \mathbf{x}_i.$$

In equilibrium, we have that

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \mathbb{E}[\theta_1 \mid \mathbf{x}_1] \\ \mathbb{E}[\theta_2 \mid \mathbf{x}_2] \\ \vdots \\ \mathbb{E}[\theta_n \mid \mathbf{x}_n] \end{bmatrix} + \begin{bmatrix} 0 \;\; \mathbb{E}[y_1 \mid \mathbf{x}_1] + w_{12}\mathbb{E}[y_2 \mid \mathbf{x}_1] + \cdots + w_{1n}\mathbb{E}[y_n \mid \mathbf{x}_1] \\ w_{21}\mathbb{E}[y_1 \mid \mathbf{x}_2] + \;\; 0 \;\; \mathbb{E}[y_2 \mid \mathbf{x}_2] + \cdots + w_{2n}\mathbb{E}[y_n \mid \mathbf{x}_2] \\ \vdots \\ w_{n1}\mathbb{E}[y_1 \mid \mathbf{x}_n] + w_{n2}\mathbb{E}[y_2 \mid \mathbf{x}_n] + \cdots + \;\; 0 \;\; \mathbb{E}[y_n \mid \mathbf{x}_n] \end{bmatrix},$$

so that

$$\begin{bmatrix} \mathbf{h}_1'\mathbf{x}_1 \\ \mathbf{h}_2'\mathbf{x}_2 \\ \vdots \\ \mathbf{h}_n'\mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}_1' \mathbf{\Lambda}' \mathbf{M}_1' \left(\mathbf{M}_1 \mathbf{M}_1'\right)^{-1} \mathbf{x}_1 \\ \boldsymbol{\phi}_2' \mathbf{\Lambda}' \mathbf{M}_2' \left(\mathbf{M}_2 \mathbf{M}_2'\right)^{-1} \mathbf{x}_2 \\ \vdots \\ \boldsymbol{\phi}_n' \mathbf{\Lambda}' \mathbf{M}_n' \left(\mathbf{M}_n \mathbf{M}_n'\right)^{-1} \mathbf{x}_n \end{bmatrix} + \begin{bmatrix} \sum_{k \neq 1} w_{1k} \mathbf{h}_k' \mathbf{M}_k \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_1' \left(\mathbf{M}_1 \mathbf{M}_1'\right)^{-1} \mathbf{x}_1 \\ \sum_{k \neq 2} w_{2k} \mathbf{h}_k' \mathbf{M}_k \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_2' \left(\mathbf{M}_2 \mathbf{M}_2'\right)^{-1} \mathbf{x}_2 \\ \vdots \\ \sum_{k \neq n} w_{nk} \mathbf{h}_k' \mathbf{M}_k \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_n' \left(\mathbf{M}_n \mathbf{M}_n'\right)^{-1} \mathbf{x}_n \end{bmatrix},$$

and, therefore, using the fact that this equation holds for any $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ we can write

$$\begin{bmatrix} \mathbf{h}_1' \mathbf{M}_1 \mathbf{M}_1' \\ \mathbf{h}_2' \mathbf{M}_2 \mathbf{M}_2' \\ \vdots \\ \mathbf{h}_n' \mathbf{M}_n \mathbf{M}_n' \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}_1' \mathbf{\Lambda}' \mathbf{M}_1' \\ \boldsymbol{\phi}_2' \mathbf{\Lambda}' \mathbf{M}_2' \\ \vdots \\ \boldsymbol{\phi}_n' \mathbf{\Lambda}' \mathbf{M}_n' \end{bmatrix} + \begin{bmatrix} \sum_{k \neq 1} w_{1k} \mathbf{h}_k' \mathbf{M}_k \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_1' \\ \sum_{k \neq 2} w_{2k} \mathbf{h}_k' \mathbf{M}_k \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_2' \\ \vdots \\ \sum_{k \neq n} w_{nk} \mathbf{h}_k' \mathbf{M}_k \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_n' \end{bmatrix}.$$

Transposing each row we get

$$\begin{bmatrix} \mathbf{M}_1\mathbf{M}_1'\mathbf{h}_1 \\ \mathbf{M}_2\mathbf{M}_2'\mathbf{h}_2 \\ \vdots \\ \mathbf{M}_n\mathbf{M}_n'\mathbf{h}_n \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1\mathbf{\Lambda}\boldsymbol{\phi}_1 \\ \mathbf{M}_2\mathbf{\Lambda}\boldsymbol{\phi}_2 \\ \vdots \\ \mathbf{M}_n\mathbf{\Lambda}\boldsymbol{\phi}_n \end{bmatrix} + \begin{bmatrix} \sum_{k\neq 1} w_{1k}\mathbf{M}_1\mathbf{\Lambda}\mathbf{\Lambda}'\mathbf{M}_k'\mathbf{h}_k \\ \sum_{k\neq 2} w_{2k}\mathbf{M}_2\mathbf{\Lambda}\mathbf{\Lambda}'\mathbf{M}_k'\mathbf{h}_k \\ \vdots \\ \sum_{k\neq n} w_{nk}\mathbf{M}_n\mathbf{\Lambda}\mathbf{\Lambda}'\mathbf{M}_k'\mathbf{h}_k \end{bmatrix},$$

which can be rewritten as

$$\overline{\mathbf{M}}\,\overline{\mathbf{M}}'\mathbf{h} = \overline{\mathbf{M}}(\mathbf{I}_n \otimes \mathbf{\Lambda})\boldsymbol{\phi} + \overline{\mathbf{M}}(\mathbf{W} \otimes \mathbf{\Lambda}\mathbf{\Lambda}')\overline{\mathbf{M}}'\mathbf{h},$$

where $\mathbf{h}$ and $\boldsymbol{\phi}$ are defined to be

$$\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \\ \vdots \\ \boldsymbol{\phi}_n \end{bmatrix}_{nu\times 1}, \quad \text{and} \quad \mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_n \end{bmatrix}.$$

Solving for $\mathbf{h}$ we obtain

$$\mathbf{h} = \mathbf{C}^{-1}\mathbf{d},$$

where

$$\begin{aligned} \mathbf{C} &\equiv \overline{\mathbf{M}}(\mathbf{I}_{nm} - \mathbf{W} \otimes \mathbf{\Lambda}\mathbf{\Lambda}')\overline{\mathbf{M}}', \\ \mathbf{d} &\equiv \overline{\mathbf{M}}(\mathbf{I}_n \otimes \mathbf{\Lambda})\boldsymbol{\phi}. \end{aligned}$$

The fact that $\mathbf{C}$ is invertible is established below in Section F4. It follows that the equilibrium to the beauty-contest model exists and is unique.

F3. Modified Signal Process and Prediction Formula

Define $\mathbf{D}$ to be

$$\mathbf{D} \equiv \sum_{i=1}^{n} \mathbf{E}_i^n \otimes \mathbf{D}_i,$$

with each $\mathbf{D}_i$ given by

$$\mathbf{D}_i \equiv \mathbf{I}_n \otimes \mathbf{\Delta} + \sum_{p=1}^{r}[\mathbf{\Omega}_p \otimes \mathbf{E}_{u+p}^m] = \operatorname{diag}\left(\tau_1^{-1/2}, \ldots, \tau_u^{-1/2}, \widetilde{\tau}_{1i}^{-1/2}, \ldots, \widetilde{\tau}_{ri}^{-1/2}\right),$$

and where $\widetilde{\tau}_{pi}^{-1}$ is the i-th eigenvalue of $(\mathbf{I}_n - \mathbf{W})^{-1}\mathbf{\Upsilon}_p$. Suppose that the signals observed by agent i are a modified version of $\mathbf{x}_i$ given by

$$\widetilde{\mathbf{x}}_i = \mathbf{M}\mathbf{D}_i\boldsymbol{\varepsilon}_i.$$

Notice that, relative to $\mathbf{\Sigma}_i$, $\mathbf{D}_i$ simply replaces the precision of the private signals τ_{pi} by the transformed $\widetilde{\tau}_{pi}$. It follows that

$$\text{(F2)} \qquad \mathbb{E}[\theta_j \mid \widetilde{\mathbf{x}}_i] = \boldsymbol{\phi}_j'\mathbf{\Lambda}'\mathbf{\Delta}\mathbf{M}'\left(\mathbf{M}\mathbf{D}_i^2\mathbf{M}'\right)^{-1}\widetilde{\mathbf{x}}_i \equiv \mathbf{g}_{ji}'\mathbf{x}_i.$$

where we used the fact that $\mathbf{\Lambda}'\mathbf{D}_i = \mathbf{\Lambda}'\mathbf{\Delta}$. Moreover, we need the following assumption.

ASSUMPTION 1: *The matrix* $(\mathbf{I}_n - \mathbf{W})$ *is invertible,* $(\mathbf{I}_n - \mathbf{W})^{-1}\mathbf{\Upsilon}_p$ *is diagonalizable, all of its* *eigenvalues have absolute value less than 1, and all of its eigenvectors are independent of p.*[4]

Under this assumption, letting $\mathbf{Q}$ denote the matrix composed of the eigenvectors of $(\mathbf{I}_n - \mathbf{W})^{-1}\boldsymbol{\Upsilon}_p$, its eigendecomposition allows us to write

$$(\mathbf{I}_n - \mathbf{W})^{-1}\boldsymbol{\Upsilon}_p = \mathbf{Q}\boldsymbol{\Omega}_p\mathbf{Q}^{-1} \tag{F3}$$

where

$$\boldsymbol{\Omega}_p \equiv \begin{bmatrix} \widetilde{\tau}_{p1}^{-1} & 0 & \cdots & 0 \\ 0 & \widetilde{\tau}_{p2}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \widetilde{\tau}_{pn}^{-1} \end{bmatrix}_{n\times n}.$$

F4. Equivalence Result

We start with a useful lemma, then, using this lemma we prove a proposition that establishes the invertibility of the matrix $\mathbf{C}$ and a final proposition that establishes the equivalence result.

LEMMA F.1: *Under Assumption 1 and using the definitions above, it follows that*

$$\boldsymbol{\Sigma}(\mathbf{I}_{nm} - \mathbf{W}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma} = [(\mathbf{I}_n - \mathbf{W})\mathbf{Q}\otimes\mathbf{I}_m]\mathbf{D}^2[\mathbf{Q}^{-1}\otimes\mathbf{I}_m]$$

PROOF:

First notice that

$$\begin{aligned}\boldsymbol{\Sigma}[(\mathbf{I}_n - \mathbf{W})\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}']\boldsymbol{\Sigma} &= \left(\sum_{i=1}^{n}\mathbf{E}_i^n\otimes\boldsymbol{\Sigma}_i\right)[(\mathbf{I}_n - \mathbf{W})\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}']\left(\sum_{j=1}^{n}\mathbf{E}_j^n\otimes\boldsymbol{\Sigma}_j\right)\\ &=\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{E}_i^n(\mathbf{I}_n - \mathbf{W})\mathbf{E}_j^n\otimes\boldsymbol{\Sigma}_i\boldsymbol{\Lambda}\boldsymbol{\Lambda}'\boldsymbol{\Sigma}_j\\ &=(\mathbf{I}_n - \mathbf{W})\otimes\boldsymbol{\Delta}^2,\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{\Sigma}[\mathbf{I}_n\otimes\boldsymbol{\Gamma}]\boldsymbol{\Sigma} &= \left(\sum_{i=1}^{n}\mathbf{E}_i^n\otimes\boldsymbol{\Sigma}_i\right)[\mathbf{I}_n\otimes\boldsymbol{\Gamma}]\left(\sum_{j=1}^{n}\mathbf{E}_j^n\otimes\boldsymbol{\Sigma}_j\right)\\ &=\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{E}_i^n\mathbf{E}_j^n\otimes\boldsymbol{\Sigma}_i\boldsymbol{\Gamma}\boldsymbol{\Sigma}_j\\ &=\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{E}_i^n\mathbf{E}_j^n\otimes\sum_{p=1}^{r}\sum_{q=1}^{r}\tau_{ip}^{-1/2}\tau_{jq}^{-1/2}\mathbf{E}_{u+p}^m\mathbf{E}_{qq}^m\\ &=\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{E}_i^n\mathbf{E}_j^n\otimes\sum_{p=1}^{r}\tau_{ip}^{-1/2}\tau_{jp}^{-1/2}\mathbf{E}_{u+p}^m\\ &=\sum_{p=1}^{r}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{E}_i^n\mathbf{E}_j^n\otimes\tau_{ip}^{-1/2}\tau_{jp}^{-1/2}\mathbf{E}_{u+p}^m\right]\end{aligned}$$

[4]A trivial case where this holds is when $\boldsymbol{\Upsilon}_p = \gamma_p\boldsymbol{\Upsilon}$ or when $r = 1$.

$$
\begin{aligned}
&= \sum_{p=1}^{r} \left[\sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{E}_i^n \tau_{ip}^{-1/2} \tau_{jp}^{-1/2} \mathbf{E}_j^n \otimes \mathbf{E}_{u+p}^m \right] \\
&= \sum_{p=1}^{r} \left[\sum_{i=1}^{n} \mathbf{E}_i^n \tau_{ip}^{-1} \otimes \mathbf{E}_{u+p}^m \right] \\
&= \sum_{p=1}^{r} \left[\boldsymbol{\Upsilon}_p \otimes \mathbf{E}_{u+p}^m \right].
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\boldsymbol{\Sigma}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma} \\
&\quad = \boldsymbol{\Sigma}(\mathbf{I}_n \otimes (\boldsymbol{\Gamma} + \boldsymbol{\Lambda}\boldsymbol{\Lambda}') - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma} \\
&\quad = \boldsymbol{\Sigma}[\mathbf{I}_n \otimes \boldsymbol{\Gamma}]\boldsymbol{\Sigma} + \boldsymbol{\Sigma}[(\mathbf{I}_n - \mathbf{W}) \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}']\boldsymbol{\Sigma} \\
&\quad = \sum_{p=1}^{r} \left[\boldsymbol{\Upsilon}_p \otimes \mathbf{E}_{u+p}^m \right] + (\mathbf{I}_n - \mathbf{W}) \otimes \boldsymbol{\Delta}^2 \\
&\quad = \sum_{p=1}^{r} \left[\boldsymbol{\Upsilon}_p \otimes \mathbf{E}_{u+p}^m \right] - (\mathbf{I}_n - \mathbf{W}) \otimes \boldsymbol{\Gamma} + (\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma}) \\
&\quad = \sum_{p=1}^{r} \left[\boldsymbol{\Upsilon}_p \otimes \mathbf{E}_{u+p}^m \right] - \sum_{p=1}^{r} \left[(\mathbf{I}_n - \mathbf{W}) \otimes \mathbf{E}_{u+p}^m \right] + (\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma}) \\
&\quad = \sum_{p=1}^{r} \left[(\boldsymbol{\Upsilon}_p - \mathbf{I}_n + \mathbf{W}) \otimes \mathbf{E}_{u+p}^m \right] + (\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma}) \\
&\quad = [(\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \left\{ \sum_{p=1}^{r} \left[(\mathbf{I}_n - \mathbf{W})^{-1} (\boldsymbol{\Upsilon}_p - \mathbf{I}_n + \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})^{-1} \mathbf{E}_{u+p}^m \right] + \mathbf{I}_{nm} \right\} \\
&\quad = [(\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \left\{ \sum_{p=1}^{r} \left[((\mathbf{I}_n - \mathbf{W})^{-1} \boldsymbol{\Upsilon}_p - \mathbf{I}_n) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})^{-1} \mathbf{E}_{u+p}^m \right] + \mathbf{I}_{nm} \right\}.
\end{aligned}
$$

Next, using equation (F3),

$$
\begin{aligned}
&\boldsymbol{\Sigma}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma} \\
&\quad = [(\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \left\{ \sum_{p=1}^{r} [(\mathbf{Q}\boldsymbol{\Omega}_p\mathbf{Q}^{-1} - \mathbf{I}_n) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})^{-1} \mathbf{E}_{u+p}^m] + \mathbf{I}_{nm} \right\} \\
&\quad = [(\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \left\{ \sum_{p=1}^{r} [(\boldsymbol{\Omega}_p - \mathbf{I}_n) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})^{-1} \mathbf{E}_{u+p}^m] + \mathbf{I}_{nm} \right\} [\mathbf{Q}^{-1} \otimes \mathbf{I}_m] \\
&\quad = [(\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes \mathbf{I}_m] \left\{ \sum_{p=1}^{r} [(\boldsymbol{\Omega}_p - \mathbf{I}_n) \otimes \mathbf{E}_{u+p}^m] + [\mathbf{I}_n \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \right\} [\mathbf{Q}^{-1} \otimes \mathbf{I}_m] \\
&\quad = [(\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes \mathbf{I}_m] \left\{ \sum_{p=1}^{r} [\boldsymbol{\Omega}_p \otimes \mathbf{E}_{u+p}^m] + (\mathbf{I}_n \otimes \boldsymbol{\Delta}^2) \right\} [\mathbf{Q}^{-1} \otimes \mathbf{I}_m]
\end{aligned}
$$

$$= [(\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes \mathbf{I}_m]\mathbf{D}^2[\mathbf{Q}^{-1} \otimes \mathbf{I}_m].$$

PROPOSITION 3: *Under Assumption 1 and using the definitions above, it follows that*

$$\begin{bmatrix} \mathbf{h}'_1 \\ \mathbf{h}'_2 \\ \vdots \\ \mathbf{h}'_n \end{bmatrix} = \sum_{k=1}^{n} \mathbf{Q}\mathbf{E}^n_{kk}\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1} \begin{bmatrix} \mathbf{g}'_{1k} \\ \mathbf{g}'_{2k} \\ \vdots \\ \mathbf{g}'_{nk} \end{bmatrix}.$$

PROOF:

First notice that

$$\begin{aligned}(\mathbf{I}_n \otimes \mathbf{M})\boldsymbol{\Sigma}(\mathbf{I}_n \otimes \boldsymbol{\Lambda}) &= (\mathbf{I}_n \otimes \mathbf{M}) \left(\sum_{i=1}^{n} \mathbf{E}^n_i \otimes \boldsymbol{\Sigma}_i \right) (\mathbf{I}_n \otimes \boldsymbol{\Lambda}) \\ &= (\mathbf{I}_n \otimes \mathbf{M}) \left(\sum_{i=1}^{n} \mathbf{E}^n_i \otimes \boldsymbol{\Sigma}_i \boldsymbol{\Lambda} \right) \\ &= (\mathbf{I}_n \otimes \mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda})\end{aligned}$$

So that, using this fact and Lemma F.1,

$$\begin{aligned}\mathbf{h} &= (\overline{\mathbf{M}}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\overline{\mathbf{M}}')^{-1}\overline{\mathbf{M}}(\mathbf{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{\phi} \\ &= \left[(\mathbf{I}_n \otimes \mathbf{M})\boldsymbol{\Sigma}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma}(\mathbf{I}_n \otimes \mathbf{M}')\right]^{-1} (\mathbf{I}_n \otimes \mathbf{M})\boldsymbol{\Sigma}(\mathbf{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{\phi} \\ &= \left[(\mathbf{I}_n \otimes \mathbf{M}) \left((\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes \mathbf{I}_m\right) \mathbf{D}^2 \left(\mathbf{Q}^{-1} \otimes \mathbf{I}_m\right) (\mathbf{I}_n \otimes \mathbf{M}')\right]^{-1} (\mathbf{I}_n \otimes \mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda})\boldsymbol{\phi} \\ &= (\mathbf{Q} \otimes \mathbf{I}_r) \left[(\mathbf{I}_n \otimes \mathbf{M})\mathbf{D}^2(\mathbf{I}_n \otimes \mathbf{M}')\right]^{-1} (\mathbf{I}_n \otimes \mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda}) \left(\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1} \otimes \mathbf{I}_u\right) \boldsymbol{\phi} \\ &= (\mathbf{Q} \otimes \mathbf{I}_r) \left(\sum_{k=1}^{n} \mathbf{E}^n_{kk} \otimes (\mathbf{M}\mathbf{D}^2_k\mathbf{M}')^{-1}\mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda} \right) \left(\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1} \otimes \mathbf{I}_u\right) \boldsymbol{\phi} \\ &= \sum_{k=1}^{n} \left[\mathbf{Q}\mathbf{E}^n_{kk}\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1} \otimes (\mathbf{M}\mathbf{D}^2_k\mathbf{M}')^{-1}\mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda}\right] \boldsymbol{\phi}.\end{aligned}$$

Finally, notice that, using the vectorization formula, we can rewrite this equation as

$$\mathbf{h} = \operatorname{vec}\left(\sum_{k=1}^{n} (\mathbf{M}\mathbf{D}^2_k\mathbf{M}')^{-1}\mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda} \begin{bmatrix} \boldsymbol{\phi}_1 & \boldsymbol{\phi}_2 & \cdots & \boldsymbol{\phi}_n \end{bmatrix} \left(\mathbf{Q}\mathbf{E}^n_k\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1}\right)' \right),$$

and, it follows from equation (F2) that

$$\mathbf{h} = \operatorname{vec}\left(\sum_{k=1}^{n} \begin{bmatrix} \mathbf{g}_{1k} & \mathbf{g}_{2k} & \cdots & \mathbf{g}_{nk} \end{bmatrix} \left(\mathbf{Q}\mathbf{E}^n_k\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1}\right)' \right).$$

The result follows by rearranging this equation.

Notice that if all agents forecast the same fundamental, i.e. $\boldsymbol{\phi}_j = \boldsymbol{\phi}_i$ then

$$\mathbf{h}_i = \sum_{k=1}^{n} \omega_{ik}\mathbf{g}_k,$$

where

$$\omega_{ik} \equiv \sum_{j=1}^{n} \mathbf{e}_i' \mathbf{Q} \mathbf{E}_k^n \mathbf{Q}^{-1} (\mathbf{I}_n - \mathbf{W})^{-1} \mathbf{e}_j,$$

and $\mathbf{e}_i$ is the i-th column of $\mathbf{I}_n$. Finally, notice that in the proof of Proposition 3 we also obtain the following corollary.

COROLLARY 1: *Under Assumption 1,* $\mathbf{C}$ *is invertible which implies that the equilibrium exists and is unique.*

Appendix G: Proof of Proposition 8

In Section G1 we obtain the canonical factorization of the auto-covariance generating function for the signal process which is necessary to apply the Wiener-Hopf prediction formula in Section G2. Section G3 presents and solves the fixed point problem that allows us to solve for the equilibrium explicitly. Section G4 describes the modified signal process and Section G5 shows the equivalence between the equilibrium policy function and the forecasting problem with the modified signal process.

G1. Canonical Factorization

This information structure is tractable enough that we can solve for the equilibrium analitically using the Wiener-Hopf prediction formula to solve the necessary forecasting problems explicitly. The observation equation is

$$\begin{bmatrix} z_t \\ x_{it} \end{bmatrix} = \underbrace{\begin{bmatrix} \tau_\varepsilon^{-1/2} & 0 & \frac{1}{1-\rho L} \\ 0 & \tau_\nu^{-1/2} & \frac{1}{1-\rho L} \end{bmatrix}}_{\equiv \mathbf{M}(L)} \underbrace{\begin{bmatrix} \hat{\varepsilon}_t \\ \hat{\nu}_{it} \\ \hat{\eta}_t \end{bmatrix}}_{\equiv \hat{\mathbf{s}}_{it}}.$$

where $\hat{\mathbf{s}}_{it}$ is a vector of standardized normal random variables. Let $\mathbf{A}(L)$ be the auto-covariance generating function for the signal process, then

$$\mathbf{A}(L) \equiv \mathbf{M}(L)\mathbf{M}'\left(L^{-1}\right) = \frac{1}{(L-\rho)(1-\rho L)} \begin{bmatrix} L + \frac{(L-\rho)(1-\rho L)}{\tau_\varepsilon} & L \\ L & L + \frac{(L-\rho)(1-\rho L)}{\tau_\nu} \end{bmatrix}.$$

In order to apply the Wiener-Hopf prediction formula we need to obtain the canonical factorization of $\mathbf{A}(L)$. Accordingly, let λ be the inside root of the determinant of $\mathbf{A}(L)$, that is

$$\lambda = \frac{1}{2}\left(\frac{\tau_\varepsilon + \tau_\nu}{\rho} + \frac{1}{\rho} + \rho - \sqrt{\left(\frac{\tau_\varepsilon + \tau_\nu}{\rho} + \frac{1}{\rho} + \rho\right)^2 - 4}\right).$$

Then notice that

$$\mathbf{V} \equiv \frac{1}{\lambda(\tau_\nu + \tau_\varepsilon)} \begin{bmatrix} \frac{\lambda\tau_\nu + \rho\tau_\varepsilon}{\tau_\varepsilon} & \rho - \lambda \\ \rho - \lambda & \frac{\rho\tau_\nu + \lambda\tau_\varepsilon}{\tau_\nu} \end{bmatrix},$$

and

$$\mathbf{B}(L) \equiv \frac{1}{(\tau_\nu + \tau_\varepsilon)(1-\rho L)} \begin{bmatrix} \tau_\nu + \tau_\varepsilon - (\rho\tau_\nu + \lambda\tau_\varepsilon) L & (\rho - \lambda)\tau_\nu L \\ (\rho-\lambda)\tau_\varepsilon L & \tau_\nu + \tau_\varepsilon - (\lambda\tau_\nu + \rho\tau_\varepsilon) L \end{bmatrix},$$

are such that

$$\mathbf{B}(L)\mathbf{V}\mathbf{B}'\left(L^{-1}\right) = \mathbf{M}(L)\mathbf{M}'\left(L^{-1}\right).$$

G2. Wiener-Hopf Prediction Formula

Applying the prediction formula, the forecast of $\theta_t = \begin{bmatrix} 0 & 0 & \frac{1}{1-\rho L} \end{bmatrix} \hat{\mathbf{s}}_{it}$ is given by

$$\begin{aligned} \mathbb{E}_{it}[\theta_t] &= \left[\begin{bmatrix} 0 & 0 & \frac{1}{1-\rho L} \end{bmatrix} \mathbf{M}'\left(L^{-1}\right)\mathbf{B}'\left(L^{-1}\right)^{-1}\right]_+ \mathbf{V}^{-1}\mathbf{B}(L)^{-1} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix} \\ &= \frac{\lambda \begin{bmatrix} \tau_\varepsilon & \tau_\nu \end{bmatrix}}{\rho(1-\lambda L)(1-\rho\lambda)} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix}. \end{aligned} \tag{G1}$$

Let $g(L) \equiv h_1(L) + h_2(L)$, then, the forecast about $y_t = \begin{bmatrix} \tau_\varepsilon^{-1/2} h_1(L) & 0 & \frac{g(L)}{1-\rho L} \end{bmatrix} \hat{\mathbf{s}}_{it}$ is given by

$$
\begin{aligned}
\mathbb{E}_{it}[y_t] =& \left[\begin{bmatrix} \tau_\varepsilon^{-1/2} h_1(L) & 0 & \frac{g(L)}{1-\rho L} \end{bmatrix} \mathbf{M}'\left(L^{-1}\right) \mathbf{B}'\left(L^{-1}\right)^{-1} \right]_+ \mathbf{V}^{-1} \mathbf{B}(L)^{-1} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix} \\
=& \left\{ \frac{\begin{bmatrix} \left(\left(\rho\tau_\nu + \lambda\tau_\varepsilon + \lambda\rho\left(\lambda\tau_\nu + \rho\tau_\varepsilon\right)\right) L - \lambda\rho\left(\tau_\nu + \tau_\varepsilon\right)\left(1 + L^2\right)\right) h_1(L) & \tau_\nu(\lambda - \rho)(1 - \rho\lambda) L h_1(L) \end{bmatrix}}{\rho\left(\tau_\nu + \tau_\varepsilon\right)(L - \lambda)(1 - \lambda L)} \right. \\
&+ \frac{\begin{bmatrix} \tau_\varepsilon(\rho - \lambda)(1 - \rho L)\lambda h_1(\lambda) & -\tau_\nu(\lambda - \rho)\lambda(1 - \rho L) h_1(\lambda) \end{bmatrix}}{\rho\left(\tau_\nu + \tau_\varepsilon\right)(L - \lambda)(1 - \lambda L)} \\
&+ \left. \frac{\lambda\left(L(1 - \rho\lambda) g(L) - \lambda(1 - \rho L) g(\lambda)\right)\begin{bmatrix} \tau_\varepsilon & \tau_\nu \end{bmatrix}}{\rho(1 - \rho\lambda)(L - \lambda)(1 - \lambda L)} \right\} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix},
\end{aligned}
$$

and the forecast about $y_{t+1} = \begin{bmatrix} \tau_\varepsilon^{-1/2} L^{-1} h_1(L) & 0 & \frac{L^{-1} g(L)}{1-\rho L} \end{bmatrix} \hat{\mathbf{s}}_{it}$ is given by

$$
\begin{aligned}
\mathbb{E}_{it}[y_{t+1}] =& \left[\begin{bmatrix} \tau_\varepsilon^{-1/2} L^{-1} h_1(L) & 0 & \frac{L^{-1} g(L)}{1-\rho L} \end{bmatrix} \mathbf{M}'\left(L^{-1}\right) \mathbf{B}'\left(L^{-1}\right)^{-1} \right]_+ \mathbf{V}^{-1} \mathbf{B}(L)^{-1} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix} \\
=& \left\{ \frac{\begin{bmatrix} \left(\left(\rho\tau_\nu + \lambda\tau_\varepsilon + \lambda\rho\left(\lambda\tau_\nu + \rho\tau_\varepsilon\right)\right) L - \lambda\rho\left(\tau_\nu + \tau_\varepsilon\right)\left(1 + L^2\right)\right) \lambda h_1(L) & \lambda\tau_\nu(\lambda - \rho)(1 - \rho\lambda) L h_1(L) \end{bmatrix}}{\rho\left(\tau_\nu + \tau_\varepsilon\right)\lambda L(L - \lambda)(1 - \lambda L)} \right. \\
&+ \frac{\begin{bmatrix} \tau_\varepsilon(\rho - \lambda)(1 - \rho L)\lambda L h_1(\lambda) & -\lambda\tau_\nu(\lambda - \rho)(1 - \rho L) L h_1(\lambda) \end{bmatrix}}{\rho\left(\tau_\nu + \tau_\varepsilon\right)\lambda L(L - \lambda)(1 - \lambda L)} \\
&+ \frac{\begin{bmatrix} -\lambda\rho(L - \lambda)\left(\tau_\nu + \tau_\varepsilon - \left(\lambda\tau_\nu + \rho\tau_\varepsilon\right) L\right) h_1(0) & \lambda\tau_\nu(\lambda - \rho) L\rho(L - \lambda) h_1(0) \end{bmatrix}}{\rho\left(\tau_\nu + \tau_\varepsilon\right)\lambda L(L - \lambda)(1 - \lambda L)} \\
&+ \left. \frac{\lambda\left((1 - \rho\lambda) g(L) - (1 - \rho L) g(\lambda)\right)\begin{bmatrix} \tau_\varepsilon & \tau_\nu \end{bmatrix}}{\rho(1 - \rho\lambda)(L - \lambda)(1 - \lambda L)} \right\} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix}.
\end{aligned}
$$

G3. Fixed Point

Substituting the forecast formulas into the best response function we obtain the following system

$$
\mathbf{C}(L) \begin{bmatrix} h_1(L) \\ h_2(L) \end{bmatrix} = \mathbf{D}(L),
$$

where

$$
\mathbf{C}(L) \equiv \begin{bmatrix} 1 - \frac{\left(\alpha + \beta L^{-1}\right)}{\rho(L-\lambda)(1-\lambda L)} \frac{\left(\rho\tau_\nu + \lambda\tau_\varepsilon + \lambda\rho\left(\lambda\tau_\nu + \rho\tau_\varepsilon\right)\right) L - \lambda\left(\tau_\nu + \tau_\varepsilon\right)\left(\rho L^2 - \tau_\varepsilon L + \rho\right)}{\left(\tau_\nu + \tau_\varepsilon\right)} & -\frac{\left(\alpha + \beta L^{-1}\right)\lambda L\tau_\varepsilon}{\rho(L-\lambda)(1-\lambda L)} \\ -\frac{\left(\alpha + \beta L^{-1}\right)}{\rho(L-\lambda)(1-\lambda L)} \frac{-\tau_\nu(\lambda-\rho)(L(\lambda\rho - 1)) + \lambda L\tau_\nu\left(\tau_\nu + \tau_\varepsilon\right)}{\left(\tau_\nu + \tau_\varepsilon\right)} & 1 - \frac{\left(\alpha + \beta L^{-1}\right)\lambda L\tau_\nu}{\rho(L-\lambda)(1-\lambda L)} \end{bmatrix},
$$

and

$$
\mathbf{D}(L) \equiv \frac{\gamma\lambda\begin{bmatrix} \tau_\varepsilon & \tau_\nu \end{bmatrix}'}{\rho(1-\lambda L)(1-\rho\lambda)} - \varphi_1 \frac{(1-\rho L)\begin{bmatrix} \tau_\varepsilon & \tau_\nu \end{bmatrix}'}{(L-\lambda)(1-\lambda L)} - \varphi_2 \frac{\begin{bmatrix} \tau_\varepsilon + \tau_\nu - \left(\lambda\tau_\nu + \rho\tau_\varepsilon\right) L & \tau_\nu(\lambda-\rho) L \end{bmatrix}'}{L(1-\lambda L)},
$$

with

$$\varphi_1 \equiv \frac{\alpha\lambda+\beta}{\lambda}\left(\frac{\lambda\left(\lambda-\rho\right)h_1\left(\lambda\right)}{\rho\left(\tau_\nu+\tau_\varepsilon\right)}+\frac{\lambda^2 g\left(\lambda\right)}{\rho\left(1-\rho\lambda\right)}\right), \quad \text{and} \quad \varphi_2 \equiv \frac{\beta}{\tau_\nu+\tau_\varepsilon}h_1\left(0\right).$$

Using the fact that $\lambda+\frac{1}{\lambda}=\rho+\frac{1}{\rho}+\frac{\tau_\varepsilon+\tau_\nu}{\rho}$ to substitute out for τ_ε, $\mathbf{C}\left(L\right)$ simplifies to

$$\mathbf{C}(L)=\begin{bmatrix}1-\alpha-\beta L^{-1} & -\frac{\alpha+\beta L^{-1}}{\rho(L-\lambda)(1-\lambda L)}\lambda L\tau_\varepsilon \\ 0 & 1-\frac{\alpha+\beta L^{-1}}{\rho(L-\lambda)(1-\lambda L)}\lambda L\tau_\nu\end{bmatrix}.$$

Next notice that the determinant of $\mathbf{C}(L)$,

$$\det(\mathbf{C}(L))=\frac{-\lambda\left(L^2-\left(\lambda+\frac{1}{\lambda}\right)L+1+\frac{\alpha L+\beta}{\rho}\tau_\nu\right)\left((1-\alpha)L-\beta\right)}{L\left(1-\lambda L\right)\left(L-\lambda\right)}$$

has two inside roots,

$$\omega_1 \equiv \frac{\rho\left(\lambda+\frac{1}{\lambda}\right)-\alpha\tau_\nu-\sqrt{\left(\rho\left(\lambda+\frac{1}{\lambda}\right)-\alpha\tau_\nu\right)^2-4\left(\rho+\beta\tau_\nu\right)\rho}}{2\rho}, \quad \text{and} \quad \omega_2 \equiv \frac{\beta}{1-\alpha},$$

and one outside root,

$$\omega_3 \equiv \frac{\rho+\beta\tau_\nu}{\rho}\frac{1}{\omega_1}.$$

Solving the system for $h_1(L)$ and $h_2(L)$ and choosing φ_1 and φ_2 to remove the inside poles of $h_1(L)$ at ω_1 and ω_2,[5] and rearranging we obtain the policy functions explicitly in terms of parameters,

$$h_1\left(L\right)=\frac{\psi}{\rho\left(\omega_3-\rho\right)}\frac{\tau_\varepsilon}{\left(1-\frac{1}{\omega_3}L\right)}, \quad \text{and} \quad h_2\left(L\right)=\frac{\psi}{\rho\left(\omega_3-\rho\right)}\frac{\left(\left(1-\alpha\right)-\frac{\beta}{\omega_3}\right)\tau_\nu}{\left(1-\frac{1}{\omega_3}L\right)}$$

where

$$\psi \equiv \frac{\gamma}{1-\alpha-\rho\beta}.$$

G4. Modified Signal Process

Suppose, now, that agents receive the same public signals as before but the private signals are given by

$$\widetilde{x}_{it}=\theta_t+\widetilde{\nu}_{it}$$

where $\widetilde{\nu}_{it}\sim\mathcal{N}(0,\widetilde{\tau}_\nu^{-1})$ and

$$\widetilde{\tau}_\nu \equiv \left(\left(1-\alpha\right)-\frac{\beta}{\omega_3}\right)\tau_\nu.$$

[5] That is,

$$\varphi_1=\frac{\gamma\rho\left(\left(1-\alpha\right)\lambda-\kappa\right)\left(1-\lambda\omega_3\right)+\gamma\kappa\lambda\tau_\nu}{\rho^2\left(1-\rho\lambda\right)\left(\omega_3-\rho\right)\left(1-\alpha-\rho\kappa\right)}, \quad \text{and} \quad \varphi_2=\frac{\gamma\kappa\left(\left(\lambda-\rho\right)\left(1-\rho\lambda\right)+\lambda\tau_\nu\right)}{\rho\left(1-\rho\lambda\right)\left(\omega_3-\rho\right)\left(1-\alpha-\rho\kappa\right)\left(\lambda-\rho\right)}.$$

Notice that, applying the Wiener-Hopf prediction formula, analogously to above, we obtain

$$\widetilde{\mathbb{E}}_{it}\left[\theta_t\right] = \frac{\widetilde{\lambda}\left[\tau_\varepsilon \quad \widetilde{\tau}_\nu\right]}{\rho\left(1-\widetilde{\lambda}L\right)\left(1-\rho\widetilde{\lambda}\right)}\begin{bmatrix} z_t \\ \widetilde{x}_{it}\end{bmatrix},$$

where

$$\widetilde{\lambda} \equiv \frac{1}{2}\left(\frac{\tau_\varepsilon + \widetilde{\tau}_\nu}{\rho} + \frac{1}{\rho} + \rho - \sqrt{\left(\frac{\tau_\varepsilon + \widetilde{\tau}_\nu}{\rho} + \frac{1}{\rho} + \rho\right)^2 - 4}\right).$$

G5. Equivalence Result

Finally, notice that

$$\widetilde{\lambda} = \omega_3^{-1},$$

and, therefore,

$$\psi\widetilde{\mathbb{E}}_{it}\left[\theta_t\right] = \frac{\psi}{\rho\left(\omega_3-\rho\right)}\frac{\left[\tau_\varepsilon \quad \left((1-\alpha)-\frac{\beta}{\omega_3}\right)\tau_\nu\right]}{\left(1-\frac{1}{\omega_3}L\right)}\begin{bmatrix} z_t \\ \widetilde{x}_{it}\end{bmatrix} = \left[h_1\left(L\right) \quad h_2\left(L\right)\right]\begin{bmatrix} z_t \\ \widetilde{x}_{it}\end{bmatrix}.$$

G6. Forward and Backward Looking Best Response

Consider the following best response function where agents also care about future and past aggregate actions,

$$y_{it} = \gamma\mathbb{E}_{it}\left[\theta_t\right] + \alpha\mathbb{E}_{it}[y_t] + \beta\mathbb{E}_{it}\left[y_{t+1}\right] + \kappa\mathbb{E}_{it}\left[y_{t-1}\right].$$

Following a similar procedure to the one above we can solve for the equilibrium best response functions,

$$h_1\left(L\right) = \frac{\psi\tau_\varepsilon}{\left(1-\frac{1}{\omega_3}L\right)\left(1-\frac{1}{\omega_4}L\right)}, \quad \text{and} \quad h_2\left(L\right) = \frac{\psi\left(\kappa\omega_3\omega_4-\beta\right)\tau_\nu}{\omega_3\left(1-\frac{1}{\omega_3}L\right)}$$

where

$$\psi \equiv \frac{\gamma\left(\rho+\kappa\tau_\nu\right)}{\left(\kappa\omega_4-\rho\beta\right)\left(\rho+\kappa\tau_\nu\right)\rho\omega_3 - \rho^2\kappa\left(\rho+\beta\tau_\nu\right)\omega_4 + \rho\beta\left(\rho^3+\left(\rho\left(1-\alpha\right)-\kappa\right)\tau_\nu\right)},$$

$$\omega_3 \equiv \frac{(1-\alpha)\,\tau_\nu + \tau_\varepsilon + 1 + \rho^2 + \sqrt{\left((1-\alpha)\,\tau_u + \tau_\varepsilon + 1 + \rho^2\right)^2 - 4\left(\rho+\beta\tau_u\right)\left(\rho+\kappa\tau_\nu\right)}}{2\left(\rho+\kappa\tau_\nu\right)},$$

$$\omega_4 \equiv \frac{1-\alpha+\sqrt{(1-\alpha)^2-4\beta\kappa}}{2\kappa}.$$

Relative to a forecast of θ_t with the signal structure above (see equation (G1)) there is an extra lag operator in the denominator of the response to the public signal, $h_1(L)$, here. Hence, the equilibrium actions cannot be represented by a forecast of θ_t by a modification of the precision of the shocks using the same information structure.

Appendix H: Additional Examples

H1. Multi-Action Example

In this section, we explore the effects of information frictions on the comovement between aggregate outcomes. To our knowledge, this paper is the first to explore this issue. The joint dynamics of multiple aggregate outcomes results from their intrinsic cross-dependence and from the degree of information frictions. Using the single-agent solution from Proposition 4, we obtain a clear characterization of these two forces. In particular, we show that increasing the degree of information frictions can flip the sign of the correlation between aggregate variables.

Consider an economy in which agents choose two actions simultaneously. For simplicity, assume that the two actions depend on the same aggregate fundamental, θ_t. Then, the agents' best response functions can be written as

$$y^1_{it} = \mathbb{E}_{it}[\theta_t] + a_{11}\mathbb{E}_{it}[y^1_t] + a_{12}\mathbb{E}_{it}[y^2_t]$$
$$y^2_{it} = \mathbb{E}_{it}[\theta_t] + a_{21}\mathbb{E}_{it}[y^1_t] + a_{22}\mathbb{E}_{it}[y^2_t]$$

where the matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \mathbf{Q}\begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix}\mathbf{Q}^{-1},$$

summarizes the dependence on aggregate actions. The second equality represents the eigendecomposition of matrix $\mathbf{A}$, where α_1 and α_2 denote its eigenvalues and $\mathbf{Q}$ is a matrix composed of its eigenvectors. Without loss of generality, let ω_1 and ω_2, be such that

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 \\ \frac{\omega_2}{\omega_1} & \frac{1-(1-\alpha_1)\omega_2}{1-(1-\alpha_1)\omega_2} \end{bmatrix}.$$

Notice that choosing α_1, α_2, ω_1 and ω_2 we can generate any matrix $\mathbf{A}$ that satisfies Assumption 4. The policy rule permits a simpler format when represented in terms of these parameters rather than the ones in the original matrix $\mathbf{A}$. Applying Proposition 4, it follows that

$$y^1_t = \omega_1\widetilde{\mathbb{E}}_t[\theta_t;\alpha_1] + \phi_1\widetilde{\mathbb{E}}_t[\theta_t;\alpha_2],$$
$$y^2_t = \omega_2\widetilde{\mathbb{E}}_t[\theta_t;\alpha_1] + \phi_2\widetilde{\mathbb{E}}_t[\theta_t;\alpha_2].$$

where ϕ_1 and ϕ_2 are functions of the primitives

$$\phi_1 \equiv \frac{1}{1-\alpha_2} - \frac{1-\alpha_1}{1-\alpha_2}\omega_1, \quad \text{and} \quad \phi_2 \equiv \frac{1}{1-\alpha_2} - \frac{1-\alpha_1}{1-\alpha_2}\omega_2.$$

This representation makes clear how the degrees of strategic complementarity affect each action. In particular, if $\alpha_1 = \alpha_2$, then $\widetilde{\mathbb{E}}_t[\theta_t;\alpha_1] = \widetilde{\mathbb{E}}_t[\theta_t;\alpha_2]$, and the actions are the same irrespective of how these forecasts are weighted. If $\alpha_1 \neq \alpha_2$, the behavior of the two actions does depend on the weights.

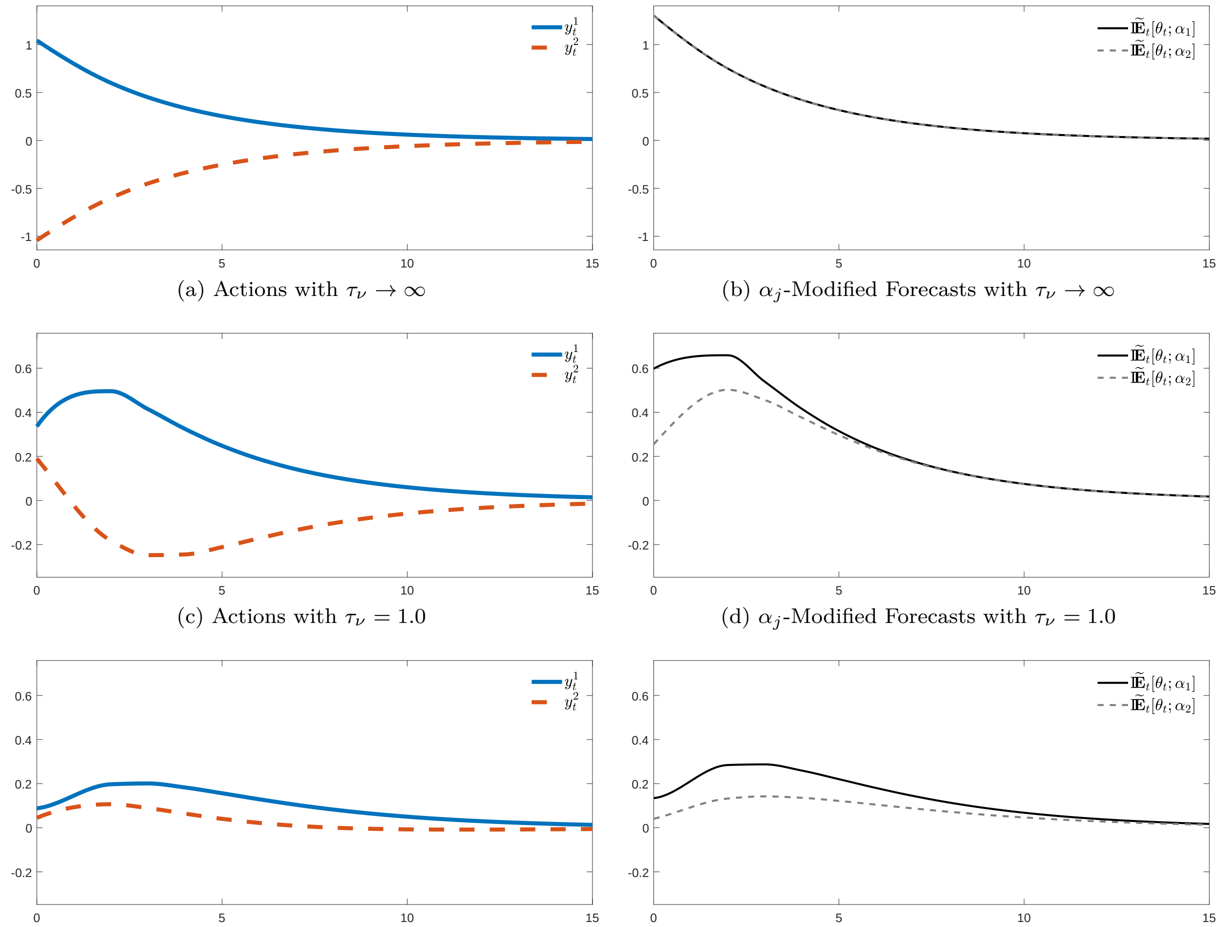
Next, we show in a numerical example how dispersed information can affect the relationship between the two actions. Suppose that the fundamental follows an AR(1) process

$$\theta_t = \rho\theta_{t-1} + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \tau_\eta^{-1}),$$

and that agents only receive a private signal about θ_t,

$$x_{it} = \theta_t + \nu_{it}, \qquad \nu_{it} \sim \mathcal{N}(0, \tau_\nu^{-1}).$$

We set the eigenvalues to $\alpha_1 = -0.5$, and $\alpha_2 = 0.5$, so the precision of the α_1-modified signals



(a) Actions with $\tau_\nu \to \infty$

(b) α_j-Modified Forecasts with $\tau_\nu \to \infty$

(c) Actions with $\tau_\nu = 1.0$

(d) α_j-Modified Forecasts with $\tau_\nu = 1.0$

(e) Actions with $\tau_\nu = 0.1$

(f) α_j-Modified Forecasts with $\tau_\nu = 0.1$

FIGURE H1. MULTI-ACTION EXAMPLE

Note: We set the following parameters, $\tau_\eta = 1$, $\rho = 0.75$, $\omega_1 = 0.6$ and $\omega_2 = 1.4$, $\alpha_1 = -0.5$ and $\alpha_2 = 0.5$.

are higher. We also set the eigenvector parameters to $\omega_1 = 0.6$, and $\omega_2 = 1.4$, which implies that $\phi_1 = 0.2$, and $\phi_2 = -2.2$. In the perfect information benchmark, both modified forecasts collapse to the fundamental itself, that is $\widetilde{\mathbb{E}}_t[\theta_t;\alpha_1] = \widetilde{\mathbb{E}}_t[\theta_t;\alpha_2] = \theta_t$. The aggregate outcomes are, then, uniquely pinned down by the fundamental and the parameters controlling strategic interactions. Explicitly,

$$y_t^1 = (\omega_1 + \phi_1)\theta_t, \quad \text{and} \quad y_t^2 = (\omega_2 + \phi_2)\theta_t.$$

Under our parameterization, $(\omega_1 + \phi_1)$ and $(\omega_2 + \phi_2)$ have opposite signs, and therefore, the actions are perfectly negatively correlated, as shown in Figure H1a. Figure H1c shows that, when $\tau_\nu = 1$, the two actions still move in different directions. The reason is that both $(1-\alpha_1)\tau_\nu$ and $(1-\alpha_2)\tau_\nu$ are large enough, so that $\widetilde{\mathbb{E}}_t[\theta_t;\alpha_1]$ and $\widetilde{\mathbb{E}}_t[\theta_t;\alpha_2]$ are still quite responsive and close to one another, as shown in Figure H1d.

When $\tau_\nu = 0.1$, the modified precisions are low enough yielding a $\widetilde{\mathbb{E}}_t[\theta_t;\alpha_2]$ close to zero; recall that $\alpha_2 = 0.5 > -0.5 = \alpha_1$. As a result, even though ϕ_1 and ϕ_2 have the opposite signs, the terms with $\widetilde{\widetilde{\mathbb{E}}}_t[\theta_t;\alpha_1]$ dominate and the aggregate actions comove. In fact, Figure H2 shows that the correlation between the two actions is decreasing in τ_ν monotonically, and switch from positive to negative.

This result—that the correlation between actions is increasing in the intensity of information frictions—depends on the particular specification of parameters in this example. However, it high-

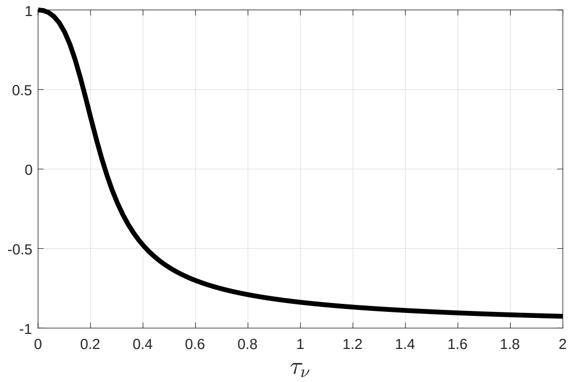


FIGURE H2. CORRELATION BETWEEN y_t^1 AND y_t^2 FOR DIFFERENT τ_ν

lights the fact that dispersed information may significantly affect the joint behavior of interactive actions.

H2. Endogenous Learning Example

Relative to Example 1, suppose agents also receive a signal about the aggregate action y_t, which is an endogenous object. Letting $\boldsymbol{\varepsilon}_{it} = [\theta_t,\, \nu_{it},\, \varepsilon_{it}]'$, the information structure can be represented within the framework of equation (25):

$$\mathbf{M}(L) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{p}(L) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \Rightarrow \qquad \begin{aligned} x_{it} &= \theta_t + \nu_{it} \\ z_{it} &= y_t + \varepsilon_{it}. \end{aligned}$$

For simplicity, we also assume that all the shocks are i.i.d.: $\theta_t \sim \mathcal{N}(0, \tau_\theta^{-1})$, $\nu_{it} \sim \mathcal{N}(0, \tau_\nu^{-1})$, and $\varepsilon_{it} \sim \mathcal{N}(0, \tau_\varepsilon^{-1})$. This type of information structure is widely used in the literature on endogenous learning.

The aggregate action, y_t, must depend on the fundamental, θ_t—the only aggregate shock in this economy—so we conjecture that

$$y_t = \mathcal{H}\,\theta_t,$$

for some constant $\mathcal{H}$. Since agents do not internalize the effects of their own actions on others' signals, they take $\mathcal{H}$ as given. Hence, their forecasting problem, given $\mathcal{H}$, is equivalent to one in which their signals are generated by the following exogenous process:

$$\widehat{\mathbf{M}}(L) = \begin{bmatrix} 1 & 1 & 0 \\ \mathcal{H} & 0 & 1 \end{bmatrix} \qquad \Rightarrow \qquad \begin{aligned} x_{it} &= \theta_t + \nu_{it} \\ \widehat{z}_{it} &= \mathcal{H}\,\theta_t + \varepsilon_{it} \end{aligned}$$

With endogenous information, even in the absence of a primitive coordinating motive ($\alpha = 0$), agents implicitly coordinate via their signal processes. Using the single-agent solution, one can focus on the alternative simple forecasting problem in which all the coordination occurs via information. In the example at hand, the aggregate action can be written as

$$y_t = \int \widetilde{\mathbb{E}}_{it}[\theta_t] = \frac{(1-\alpha)\tau_\nu}{\tau_\theta + (1-\alpha)\tau_\nu + \mathcal{H}^2\,(1-\alpha)\tau_\varepsilon} \int x_{it} + \frac{\mathcal{H}\,(1-\alpha)\tau_\varepsilon}{\tau_\theta + (1-\alpha)\tau_\nu + \mathcal{H}^2\,(1-\alpha)\tau_\varepsilon} \int \widehat{z}_{it}.$$

At this stage, the equivalence result spares us the trouble of making an inference about y_t, and this policy rule already satisfies the first two equilibrium conditions in Definition 3. To make sure that the perceived law of motion for y_t is consistent with agents' signal processes, condition (27) must also be satisfied, which reduces to a cubic polynomial equation in terms of $\mathcal{H}$,

$$(1-\alpha)\tau_\varepsilon\,\mathcal{H}^3 - (1-\alpha)\tau_\varepsilon\,\mathcal{H}^2 + (\tau_\theta + (1-\alpha)\tau_\nu)\,\mathcal{H} - (1-\alpha)\tau_\nu = 0. \tag{H1}$$

It is clear that there may exist multiple real solutions to equation (H1), which correspond to multiple equilibria. The origin of this multiplicity lies in the self-fulfilling property of the signals' informativeness. For example, if all agents respond to the fundamental aggressively, if $\mathcal{H}$ is high, then the signal z_{it} is very informative. As a result, agents can learn more from the endogenous signal and indeed become more responsive.

*

REFERENCES

Angeletos, George-Marios, and Jennifer La'O. 2010. "Noisy Business Cycles." In *NBER Macroeconomics Annual 2009, Volume 24*. 319–378. University of Chicago Press.

Hamilton, James Douglas. 1994. *Time Series Analysis.* . 2 ed., Princeton University Press.

Townsend, Robert M. 1983. "Forecasting the Forecasts of Others." *Journal of Political Economy*, 91(4): 546–88.

Whittle, Peter. 1963. *Prediction and Regulation by Linear Least-Square Methods.* University of Minnesota Press.

Woodford, Michael. 2002. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honour of Edmund S. Phelps.* , ed. P. Aghion, R. Frydman, J. Stiglitz and M.Woodford. Princeton University Press.