BANK OF ENGLAND

Appendix to Working Paper No. 503

Peering into the mist: social learning over an opaque observation network

John Barrdear

August 2014

1 Proof of Lemma 1.

I here demonstrate that agents' contemporaneous expectations of the network shocks are zero:

$$E_t\left(i\right)\left[\boldsymbol{z}_t\right] = \mathbf{0}\ \forall i, t \tag{1}$$

Since all shocks are Gaussian, the ability of an agent to create an expectation about a variable depends on the covariance between that variable and the agent's signal vector. But, by construction, agent i does not observe any signal that is based on $\boldsymbol{z}_t$. Since $\boldsymbol{z}_t$ is transitory and fully independent across time and from the underlying state, it must be the case that $Cov\left(\boldsymbol{z}_t, \boldsymbol{s}_t\left(i\right)\right) = \mathbf{0}$. The only possible exception to this is to note that $\boldsymbol{z}_t$ is comprised of weighted sums of idiosyncratic shocks and agent i's signals do include $\boldsymbol{v}_t\left(i\right)$. However, it must be that:

$$\begin{aligned} Cov\left({}^{1:}\widetilde{\boldsymbol{v}}_t, \boldsymbol{v}_t\left(i\right)\right) &= E\left[\lim_{n\to\infty}\sum_{j=1}^{n}\phi_n\left(j\right)\boldsymbol{v}_t\left(j\right)\boldsymbol{v}_t\left(i\right)\right] \\ &= \lim_{n\to\infty}\phi_n\left(i\right)\Sigma_{vv} \\ &= \mathbf{0} \end{aligned}$$

where the second equality relies on the independence of agents' idiosyncratic shocks and the third on assumption 2 (which grants us that $\lim_{n\to\infty}\phi_n\left(i\right) = 0\ \forall i$). An equivalent argument applies to all higher-weighted averages: $Cov\left({}^{p:}\widetilde{\boldsymbol{v}}_t, \boldsymbol{v}_t\left(j\right)\right)$.

2 Proof of lemma 2.

The Kalman filter requires that each agent construct a prior expectation of the signal she will receive and then update her beliefs on the basis of the extent to which the signal she actually receives is a surprise. Using the equation for each agent's decision rule, we have that when preparing for period $t+1$, agent i will construct her prior expectation of her social signal as follows:

$$E_t\left(i\right)\left[g_t\left(\delta_t\left(i\right)\right)\right] = E_t\left(i\right)\left[\boldsymbol{\lambda}_1' E_t\left(\delta_t\left(i\right)\right)\left[X_t\right] + \boldsymbol{\lambda}_2'\boldsymbol{x}_t + \boldsymbol{\lambda}_3'\boldsymbol{v}_t\left(\delta_t\left(i\right)\right)\right]$$

Recall that $\delta_t\left(i\right)$ is not known to agent i until period $t+1$. By denying agents knowledge of the full network and, instead, granting them knowledge of the distribution from which observation links are drawn ($\boldsymbol{\Phi}$) and using the assumption that this distribution is independent of other shocks, we can note that:

$$\begin{aligned} E_t\left(i\right)\left[g_t\left(\delta_t\left(i\right)\right)\right] &= \int_0^1 E_t\left(i\right)\left[g_t\left(j\right)\right]\phi\left(j\right)dj \\ &= E_t\left(i\right)\left[\int_0^1 g_t\left(j\right)\phi\left(j\right)dj\right] \\ &= E_t\left(i\right)\left[\widetilde{g}_t\right] \\ &= E_t\left(i\right)\left[\boldsymbol{\lambda}_1'\widetilde{E}_t\left[X_t\right] + \boldsymbol{\lambda}_2'\boldsymbol{x}_t + \boldsymbol{\lambda}_3'\widetilde{\boldsymbol{v}}_t\right] \end{aligned}$$

where the second equality exploits the linearity of the expectation operator. The object $\widetilde{g}_t \equiv \int_0^1 g_t\left(j\right)\phi\left(j\right)dj$ is a *weighted* average of all agents' actions in period t using the observation p.d.f. as the weights.

3 Proof of proposition 1.

Denoting $\zeta(n) \equiv \sum_{i=1}^{n} \phi_n(i)^2$ and assuming that $\lim_{n\to\infty} \zeta(n) = \zeta^* \in (0,1)$ (assumption 2), we here demonstrate the following results regarding agents' idiosyncratic shocks:

1. ${}^{p:}\tilde{\boldsymbol{v}}_{n,t} \xrightarrow{d} {}^{p:}\tilde{\boldsymbol{v}}_t \; \forall p$ where ${}^{p:}\tilde{\boldsymbol{v}}_t \sim N\left(\mathbf{0}, \Sigma^{\{p\}}_{\underset{\sim}{vv}}\right) \quad \Sigma^{\{q\}}_{\underset{\sim}{vv}} = (1-(1-\zeta^*)^q)\,\Sigma_{vv}$
2. ${}^{p:}\ddot{\boldsymbol{v}}_{n,t} \xrightarrow{\mathcal{L}^2} {}^{p:}\tilde{\boldsymbol{v}}_t \; \forall p$
3. $Cov\left({}^{p:}\tilde{\boldsymbol{v}}_t, {}^{r:}\tilde{\boldsymbol{v}}_t\right) = \Sigma^{\{p\}}_{\underset{\sim}{vv}} \; \forall r < q$

where the weighted sums are defined as:

$$
\begin{aligned}
{}^{1:}\tilde{\boldsymbol{v}}_{n,t} &\equiv \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{v}_t(\delta_t(i)) & {}^{1:}\ddot{\boldsymbol{v}}_{n,t} &\equiv \sum_{i=1}^{n} \boldsymbol{v}_t(i)\,\phi_n(i) \\
{}^{2:}\tilde{\boldsymbol{v}}_{n,t} &\equiv \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{v}_t(\delta_t(\delta_t(i))) & {}^{2:}\ddot{\boldsymbol{v}}_{n,t} &\equiv \sum_{i=1}^{n} \boldsymbol{v}_t(\delta_t(i))\,\phi_n(i) \\
{}^{3:}\tilde{\boldsymbol{v}}_{n,t} &\equiv \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{v}_t(\delta_t(\delta_t(\delta_t(i)))) & {}^{3:}\ddot{\boldsymbol{v}}_{n,t} &\equiv \sum_{i=1}^{n} \boldsymbol{v}_t(\delta_t(\delta_t(i)))\,\phi_n(i) \\
&\vdots & &\vdots
\end{aligned}
$$

First, note that since the vector $\boldsymbol{v}_t(i)$ is drawn from independent and identical Gaussian distributions with mean zero for each i and t, all of the weighted sums must also be distributed normally with mean zero. We now consider each of the results in turn.

1. ${}^{p:}\tilde{\boldsymbol{v}}_{n,t} \xrightarrow{d} {}^{p:}\tilde{\boldsymbol{v}}_t \; \forall p \quad {}^{p:}\tilde{\boldsymbol{v}}_t \sim N\left(\mathbf{0}, \Sigma^{\{p\}}_{\underset{\sim}{vv}}\right) \quad \Sigma^{\{p\}}_{\underset{\sim}{vv}} = (1-(1-\zeta^*)^p)\,\Sigma_{vv}$

Since it is clear that ${}^{p:}\tilde{\boldsymbol{v}}_{n,t}$ must converge to a normal distribution with mean zero, all that remains is to determine its variance-covariance matrix (note that the law of large numbers will apply here when the variance-covariance matrix is zero).

We will begin by considering each weighted-sum in turn.

- ${}^{1:}\tilde{\boldsymbol{v}}_{n,t} \xrightarrow{d} {}^{1:}\tilde{\boldsymbol{v}}_{t}$

The variance of ${}^{1:}\tilde{\boldsymbol{v}}_{n,t}$ is given by:

$$
\begin{aligned}
Var\left[{}^{1:}\tilde{\boldsymbol{v}}_{n,t}\right] &= \frac{1}{n^2} Var\left[\boldsymbol{v}_t\left(\delta_t\left(1\right)\right) + \boldsymbol{v}_t\left(\delta_t\left(2\right)\right) + \cdots + \boldsymbol{v}_t\left(\delta_t\left(n\right)\right)\right] \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left[\boldsymbol{v}_t\left(\delta_t\left(i\right)\right) \boldsymbol{v}_t\left(\delta_t\left(j\right)\right)\right] \\
&= \frac{1}{n^2} \left(n\Sigma_{vv} + \sum_{i=1}^{n} \sum_{j\neq i}^{n} E\left[\boldsymbol{v}_t\left(\delta_t\left(i\right)\right) \boldsymbol{v}_t\left(\delta_t\left(j\right)\right)\right] \right)
\end{aligned}
$$

However, when $i \neq j$, given the full independence of the distributions of agents' observees, it must be that

$$
\begin{aligned}
E\left[\boldsymbol{v}_t\left(\delta_t\left(i\right)\right) \boldsymbol{v}_t\left(\delta_t\left(j\right)\right)\right] &= \sum_{k=1}^{n} \phi_n\left(k\right) E\left[\boldsymbol{v}_t\left(k\right) \boldsymbol{v}_t\left(\delta_t\left(j\right)\right)\right] \\
&= \sum_{k=1}^{n} \phi_n\left(k\right) \left(\sum_{l=1}^{n} \phi_n\left(l\right) E\left[\boldsymbol{v}_t\left(k\right) \boldsymbol{v}_t\left(l\right)\right] \right) \\
&= \sum_{k=1}^{n} \phi_n\left(k\right)^2 E\left[\boldsymbol{v}_t\left(k\right) \boldsymbol{v}_t\left(k\right)\right] \\
&= \zeta\left(n\right) \Sigma_{vv} \qquad (2)
\end{aligned}
$$

where in moving from the second line to the third we have made use of the independence of agents' idiosyncratic shocks. We therefore have that

$$
\begin{aligned}
Var\left[{}^{1:}\tilde{\boldsymbol{v}}_{n,t}\right] &= \frac{1}{n^2} \left(n\Sigma_{vv} + \sum_{i=1}^{n} \sum_{j\neq i}^{n} \zeta\left(n\right) \Sigma_{vv} \right) \\
&= \frac{1}{n^2} \left(n\Sigma_{vv} + \left(n^2 - n\right) \zeta\left(n\right) \Sigma_{vv} \right) \\
&= \frac{1}{n} \Sigma_{vv} + \left(\frac{n-1}{n} \right) \zeta\left(n\right) \Sigma_{vv}
\end{aligned}
$$

and thus, in the limit, it must be that

$$
\Sigma^{\{1\}}_{\widetilde{vv}} \equiv \lim_{n\to\infty} Var\left[{}^{1:}\tilde{\boldsymbol{v}}_{n,t}\right] = \zeta^{*} \Sigma_{vv} \qquad (3)
$$

- ${}^{2:}\tilde{\boldsymbol{v}}_{n,t} \xrightarrow{d} {}^{2:}\tilde{\boldsymbol{v}}_{t}$

The variance of ${}^{2:}\tilde{\boldsymbol{v}}_{n,t}$ is given by:

$$\begin{aligned} Var\left[{}^{2:}\tilde{\boldsymbol{v}}_{n,t}\right] &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(i\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right] \\ &= \frac{1}{n^2}\left(n\Sigma_{vv} + \sum_{i=1}^{n}\sum_{j\neq i}^{n} E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(i\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right]\right) \end{aligned}$$

Focussing on the latter term, we have that when $i \neq j$, it must be that

$$\begin{aligned} E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(i\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right] &= \sum_{k=1}^{n}\phi_n\left(k\right)E\left[\boldsymbol{v}_t\left(\delta_t\left(k\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right] \\ &= \sum_{k=1}^{n}\phi_n\left(k\right)\left(\sum_{l=1}^{n}\phi_n\left(l\right)E\left[\boldsymbol{v}_t\left(\delta_t\left(k\right)\right)\boldsymbol{v}_t\left(\delta_t\left(l\right)\right)\right]\right) \\ &= \sum_{k=1}^{n}\phi_n\left(k\right)^2\Sigma_{vv} \\ &\quad + \sum_{k=1}^{n}\sum_{l\neq k}^{n}\phi_n\left(k\right)\phi_n\left(l\right)E\left[\boldsymbol{v}_t\left(\delta_t\left(k\right)\right)\boldsymbol{v}_t\left(\delta_t\left(l\right)\right)\right] \end{aligned}$$

It was shown above in equation (2) that

$$E\left[\boldsymbol{v}_t\left(\delta_t\left(k\right)\right)\boldsymbol{v}_t\left(\delta_t\left(l\right)\right)\right] = \zeta\left(n\right)\Sigma_{vv}\ \forall k \neq l$$

so it follows that

$$E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(i\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right] = \zeta\left(n\right)\Sigma_{vv} + \zeta\left(n\right)\Sigma_{vv}\sum_{k=1}^{n}\sum_{l\neq k}^{n}\phi_n\left(k\right)\phi_n\left(l\right)$$

next, consider that since $\phi_n\left(k\right)$ and $\phi_n\left(l\right)$ are p.d.fs, it must be that

$$\begin{aligned} \sum_{k=1}^{n}\sum_{l=1}^{n}\phi_n\left(i\right)\phi_n\left(j\right) &= \sum_{k=1}^{n}\phi_n\left(k\right)\left(\sum_{l=1}^{n}\phi_n\left(l\right)\right) \\ &= \sum_{k=1}^{n}\phi_n\left(k\right) \\ &= 1 \end{aligned}$$

We must therefore have that

$$\sum_{k=1}^{n}\sum_{l\neq k}^{n}\phi_n\left(k\right)\phi_n\left(l\right) = 1 - \sum_{k=1}^{n}\phi_n\left(k\right)^2 = 1 - \zeta\left(n\right) \tag{4}$$

Thus, when $i \neq j$, we have

$$E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(i\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right] = \zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\zeta\left(n\right)\Sigma_{vv} \tag{5}$$

Substituting this back in, we arrive at

$$\begin{aligned} Var\left[{}^{2:}\tilde{\boldsymbol{v}}_{n,t}\right] &= \frac{1}{n}\Sigma_{vv} + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\left(\zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\zeta\left(n\right)\Sigma_{vv}\right) \\ &= \frac{1}{n}\Sigma_{vv} + \frac{n\left(n-1\right)}{n^2}\left(\zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\zeta\left(n\right)\Sigma_{vv}\right) \end{aligned}$$

and thus, in the limit, it must be that

$$\Sigma^{\{2\}}_{\underset{\sim}{\tilde{v}\tilde{v}}} \equiv \lim_{n\rightarrow\infty} Var\left[{}^{2:}\tilde{\boldsymbol{v}}_{n,t}\right] = \zeta^*\Sigma_{vv} + \left(1-\zeta^*\right)\zeta^*\Sigma_{vv} \tag{6}$$

- ${}^{3:\sim}\boldsymbol{v}_{n,t} \xrightarrow{d} {}^{3:\sim}\boldsymbol{v}_t$

The variance of ${}^{3:\sim}\boldsymbol{v}_{n,t}$ is given by:

$$
\begin{aligned}
Var\left[{}^{3:\sim}\boldsymbol{v}_{n,t}\right] &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(\delta_t\left(i\right)\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right)\right] \\
&= \frac{1}{n^2}\left(n\Sigma_{vv} + \sum_{i=1}^{n}\sum_{j\neq i}^{n} E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(\delta_t\left(i\right)\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right)\right]\right)
\end{aligned}
$$

Focussing on the latter term, we have that when $i \neq j$, it must be that

$$
\begin{aligned}
&E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(\delta_t\left(i\right)\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right)\right] \\
&= \sum_{k=1}^{n}\phi_n\left(k\right)\left(\sum_{l=1}^{n}\phi_n\left(l\right)E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(k\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(l\right)\right)\right)\right]\right) \\
&= \sum_{k=1}^{n}\phi_n\left(k\right)^2\Sigma_{vv} + \sum_{k=1}^{n}\sum_{l\neq k}^{n}\phi_n\left(k\right)\phi_n\left(l\right)E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(k\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(l\right)\right)\right)\right]
\end{aligned}
$$

It was shown above in equation (5) that

$$
E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(k\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(l\right)\right)\right)\right] = \zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\zeta\left(n\right)\Sigma_{vv}
$$

Combined with equation (4), this then implies that when $i \neq j$,

$$
\begin{aligned}
&E\left[\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(\delta_t\left(i\right)\right)\right)\right)\boldsymbol{v}_t\left(\delta_t\left(\delta_t\left(\delta_t\left(j\right)\right)\right)\right)\right] \\
&= \zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\left(\zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\zeta\left(n\right)\Sigma_{vv}\right)
\end{aligned}
\tag{7}
$$

Substituting this back in, we arrive at

$$
\begin{aligned}
Var\left[{}^{3:\sim}\boldsymbol{v}_{n,t}\right] &= \frac{1}{n}\Sigma_{vv} \\
&\quad + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\left(\zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\left(\zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\zeta\left(n\right)\Sigma_{vv}\right)\right) \\
&= \frac{1}{n}\Sigma_{vv} \\
&\quad + \frac{n\left(n-1\right)}{n^2}\left(\zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\left(\zeta\left(n\right)\Sigma_{vv} + \left(1-\zeta\left(n\right)\right)\zeta\left(n\right)\Sigma_{vv}\right)\right)
\end{aligned}
$$

and thus, in the limit, it must be that

$$
\Sigma_{\widetilde{vv}}^{\{3\}} \equiv \lim_{n\to\infty} Var\left[{}^{3:\sim}\boldsymbol{v}_{n,t}\right] = \zeta^*\Sigma_{vv} + \left(1-\zeta^*\right)\left(\zeta^*\Sigma_{vv} + \left(1-\zeta^*\right)\zeta^*\Sigma_{vv}\right) \tag{8}
$$

- The general case

By this stage, it should be clear that the variance-covariance matricies of higher weighted averages of agents' idiosyncratic shocks are able to be expressed in a recursive form:

$$\Sigma_{\widetilde{vv}}^{\{q\}} = \zeta^{*}\Sigma_{vv} + (1-\zeta^{*})\,\Sigma_{\widetilde{vv}}^{\{q-1\}}$$

This may be simplified by first expanding it as

$$\begin{aligned}
\Sigma_{\widetilde{vv}}^{\{q\}} &= \left(\sum_{p=0}^{q-1}(1-\zeta^{*})^{p}\right)\zeta^{*}\Sigma_{vv} \\
&= \left(\frac{1-(1-\zeta^{*})^{q}}{1-(1-\zeta^{*})}\right)\zeta^{*}\Sigma_{vv} \\
&= \left(1-(1-\zeta^{*})^{q}\right)\Sigma_{vv} \qquad (9)
\end{aligned}$$

which completes the proof of the first result.

As a matter of curiosity, this result also obtains from the following when taking each variable in $\boldsymbol{v}_t\,(i)$ separately (for simplicity I have shown only three agents, when there are actually $n \to \infty$):

$$\Sigma_{\widetilde{vv}}^{\{1\}} = \phi' \begin{bmatrix} \sigma_v^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{bmatrix} \phi = \phi'\phi\,\sigma_v^2$$

$$\begin{aligned}
\Sigma_{\widetilde{vv}}^{\{2\}} &= \phi' \begin{bmatrix} \sigma_v^2 & \Sigma_{\widetilde{vv}}^{\{1\}} & \Sigma_{\widetilde{vv}}^{\{1\}} \\ \Sigma_{\widetilde{vv}}^{\{1\}} & \sigma_v^2 & \Sigma_{\widetilde{vv}}^{\{1\}} \\ \Sigma_{\widetilde{vv}}^{\{1\}} & \Sigma_{\widetilde{vv}}^{\{1\}} & \sigma_v^2 \end{bmatrix} \phi = \phi' \begin{bmatrix} \sigma_v^2 & \phi' \begin{bmatrix} \sigma_v^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{bmatrix} \phi & \phi' \begin{bmatrix} \sigma_v^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{bmatrix} \phi \\ \phi' \begin{bmatrix} \sigma_v^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{bmatrix} \phi & \sigma_v^2 & \phi' \begin{bmatrix} \sigma_v^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{bmatrix} \phi \\ \phi' \begin{bmatrix} \sigma_v^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{bmatrix} \phi & \phi' \begin{bmatrix} \sigma_v^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{bmatrix} \phi & \sigma_v^2 \end{bmatrix} \phi \\
&= \phi' \begin{bmatrix} 1 & \phi'\phi & \phi'\phi \\ \phi'\phi & 1 & \phi'\phi \\ \phi'\phi & \phi'\phi & 1 \end{bmatrix} \phi\,\sigma_v^2 \\
&= \phi'\phi\left(1+(1-\phi'\phi)\right)\sigma_v^2
\end{aligned}$$

$$\Sigma_{\widetilde{vv}}^{\{p\}} = \phi' \begin{bmatrix} \sigma_v^2 & \Sigma_{\widetilde{vv}}^{\{p-1\}} & \Sigma_{\widetilde{vv}}^{\{p-1\}} \\ \Sigma_{\widetilde{vv}}^{\{p-1\}} & \sigma_v^2 & \Sigma_{\widetilde{vv}}^{\{p-1\}} \\ \Sigma_{\widetilde{vv}}^{\{p-1\}} & \Sigma_{\widetilde{vv}}^{\{p-1\}} & \sigma_v^2 \end{bmatrix} \phi = \phi'\phi\left(1+(1-\phi'\phi)+\cdots+(1-\phi'\phi)^{p-1}\right)\sigma_v^2$$

2. $\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t} \xrightarrow{\mathcal{L}^2} \overset{p:\sim}{\boldsymbol{v}}_t \; \forall q$

We next demonstrate that $\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}$ converges to $\overset{p:\sim}{\boldsymbol{v}}_t$ in mean square error.[1] That is, we show that $\lim_{n\to\infty} E\left[\left(\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t} - \overset{p:\sim}{\boldsymbol{v}}_t\right)^2\right] = 0$. First, see that:

$$\begin{aligned} E\left[\left(\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t} - \overset{2:\sim}{\boldsymbol{v}}_t\right)^2\right] &= E\left[\left(\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}\right)^2 - 2\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}\tilde{\boldsymbol{v}}_t + \left(\overset{p:\sim}{\boldsymbol{v}}_t\right)^2\right] \\ &= Var\left[\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}\right] - 2Cov\left[\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}, \overset{p:\sim}{\boldsymbol{v}}_t\right] + Var\left[\overset{p:\sim}{\boldsymbol{v}}_t\right] \end{aligned}$$

The third term is just $\Sigma_{\widetilde{vv}}^{\{q\}}$ from the first result above. We now consider the first and second terms in turn. The variance of $\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}$ is given by:

$$\begin{aligned} Var\left[\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}\right] &= Var\left[\sum_{i=1}^{n} \phi_n(i)\, \boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}(i)))\right)\right] \\ &= E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} \phi_n(i)\,\phi_n(j)\, \boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}(i)))\right) \boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}(j)))\right)\right] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \phi_n(i)\,\phi_n(j)\, E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}(i)))\right) \boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}(j)))\right)\right] \\ &= \sum_{i=1}^{n} \phi_n(i)^2\, \Sigma_{vv} \\ &\quad + \sum_{i=1}^{n}\sum_{j\neq i}^{n} \phi_n(i)\,\phi_n(j)\, E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}(i)))\right) \boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}(j)))\right)\right] \end{aligned}$$

But we know from the first result above that when $i \neq j$,

$$\begin{aligned} & E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}(i)))\right) \boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}(j)))\right)\right] \\ &= \zeta(n)\,\Sigma_{vv} + (1-\zeta(n))\, E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-2}(i)))\right) \boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-2}(j)))\right)\right] \end{aligned}$$

noting the recursive structure and making use of equation (4) then gives us

$$Var\left[\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}\right] = \zeta(n)\,\Sigma_{vv} + (1-\zeta(n))\, Var\left[\overset{p-1:\cdot\cdot}{\boldsymbol{v}}_{n,t}\right]$$

which, in the limit, becomes

$$\lim_{n\to\infty} Var\left[\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}\right] = \zeta^*\Sigma_{vv} + (1-\zeta^*) \lim_{n\to\infty} Var\left[\overset{p-1:\cdot\cdot}{\boldsymbol{v}}_{n,t}\right]$$

which is the same rule for $Var\left[\overset{p:\sim}{\boldsymbol{v}}_{n,t}\right]$, which implies that

$$\lim_{n\to\infty} Var\left[\overset{p:\cdot\cdot}{\boldsymbol{v}}_{n,t}\right] = \lim_{n\to\infty} Var\left[\overset{p:\sim}{\boldsymbol{v}}_{n,t}\right] = \Sigma_{\widetilde{vv}}^{\{p\}}$$

[1] Recall that convergence in mean square error is a stronger form of convergence than convergence in probability.

Turning next to the covariance between ${}^{p:}\ddot{\boldsymbol{v}}_{n,t}$ and ${}^{p:}\widetilde{\boldsymbol{v}}_t$, we note that

$$
\begin{aligned}
Cov\left[{}^{p:}\ddot{\boldsymbol{v}}_{n,t}, {}^{p:}\widetilde{\boldsymbol{v}}_{n,t}\right] &= E\left[\begin{array}{c}\left(\sum_{i=1}^{n}\phi_n\left(i\right)\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}\left(i\right)))\right)\right) \\ \times\left(\frac{1}{n}\sum_{j=1}^{n}\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p}\left(j\right)))\right)\right)\end{array}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\phi_n\left(i\right)E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}\left(i\right)))\right)\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p}\left(j\right)))\right)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\phi_n\left(i\right)\phi_n\left(k\right)E\left[\begin{array}{c}\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}\left(i\right)))\right) \\ \times\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}\left(k\right)))\right)\end{array}\right]
\end{aligned}
$$

where moving from the second line to the third makes use of the independence of agents' draws from Φ_n and the linearity of the expectation operator. This, in turn, may be rewritten as

$$
Cov\left[{}^{p:}\ddot{\boldsymbol{v}}_{n,t}, {}^{p:}\widetilde{\boldsymbol{v}}_{n,t}\right] = \frac{n}{n}\left(\begin{array}{c}\sum_{i=1}^{n}\phi_n\left(i\right)^2\Sigma_{vv} \\ \\ +\sum_{i=1}^{n}\sum_{k\neq i}^{n}\phi_n\left(i\right)\phi_n\left(k\right)E\left[\begin{array}{c}\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}\left(i\right)))\right) \\ \times\boldsymbol{v}_t\left(\underbrace{\delta_t(\cdots(\delta_t}_{p-1}\left(k\right)))\right)\end{array}\right]\end{array}\right)
$$

Since this is the same expression as that for $Var\left[{}^{p:}\ddot{\boldsymbol{v}}_{n,t}\right]$ above, we therefore have

$$
\lim_{n\to\infty}Cov\left[{}^{p:}\ddot{\boldsymbol{v}}_{n,t}, {}^{p:}\widetilde{\boldsymbol{v}}_{n,t}\right] = \Sigma^{\{p\}}_{\underset{\sim}{vv}}
$$

and, hence, that

$$
\begin{aligned}
\lim_{n\to\infty}E\left[\left({}^{p:}\ddot{\boldsymbol{v}}_{n,t} - {}^{p:}\widetilde{\boldsymbol{v}}_t\right)^2\right] &= \Sigma^{\{p\}}_{\underset{\sim}{vv}} - 2\Sigma^{\{p\}}_{\underset{\sim}{vv}} + \Sigma^{\{p\}}_{\underset{\sim}{vv}} \\
&= 0
\end{aligned}
$$

as required.

3. $Cov\left[{}^{p:}\widetilde{\boldsymbol{v}}_t, {}^{r:}\widetilde{\boldsymbol{v}}_t\right] = \Sigma^{\{p\}}_{\underset{\sim}{vv}}\ \forall p < r$

To prove this, we will first consider $Cov\left[{}^{p:}\widetilde{\boldsymbol{v}}_t, {}^{p+1:}\widetilde{\boldsymbol{v}}_t\right]$ and later consider $r \geq p+2$.

$$
\begin{aligned}
Cov\left[{}^{\{p\}}\widetilde{\boldsymbol{v}}_{n,t}, {}^{\{p+1\}}\widetilde{\boldsymbol{v}}_{n,t}\right] &= E\left[\begin{array}{c}\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}\left(i\right)\right)\right) \\ \times\left(\frac{1}{n}\sum_{j=1}^{n}\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p+1}\left(j\right)\right)\right)\end{array}\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}\left(i\right)\right)\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p+1}\left(j\right)\right)\right]
\end{aligned}
$$

Focussing on the final term, note that

$$
\begin{aligned}
&E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(i)\right)\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p+1}(j)\right)\right]\\
&=\sum_{k=1}^{n}\phi_n(k)\,E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(i)\right)\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(k)\right)\right]\\
&=\phi_n(i)\,\Sigma_{vv}+\sum_{k\neq i}^{n}\phi_n(k)\,E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(i)\right)\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(k)\right)\right]\\
&=\phi_n(i)\,\Sigma_{vv}+(1-\phi_n(i))\,\Sigma^{p}_{\underset{\sim}{vv}}(n)
\end{aligned}
$$

Substituting this back into the above then gives

$$
\begin{aligned}
Cov\left[{}^{p:\sim}\boldsymbol{v}_t,\ {}^{p+1:\sim}\boldsymbol{v}_t\right]&=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\phi_n(i)\,\Sigma_{vv}+(1-\phi_n(i))\,\Sigma^{p}_{\underset{\sim}{vv}}(n)\right)\\
&=\frac{1}{n}\sum_{i=1}^{n}\left(\phi_n(i)\,\Sigma_{vv}+(1-\phi_n(i))\,\Sigma^{p}_{\underset{\sim}{vv}}(n)\right)\\
&=\frac{1}{n}\Sigma_{vv}+\frac{1}{n}\sum_{i=1}^{n}(1-\phi_n(i))\,\Sigma^{p}_{\underset{\sim}{vv}}(n)
\end{aligned}
$$

In the limit, this becomes

$$
\lim_{n\to\infty} Cov\left[{}^{p:\sim}\boldsymbol{v}_{n,t},\ {}^{p+1:\sim}\boldsymbol{v}_{n,t}\right]=\Sigma^{p}_{\underset{\sim}{vv}}
$$

which establishes the result for $r=p+1$. For $r=p+2$, note that

$$
\begin{aligned}
&E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(i)\right)\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p+2}(j)\right)\right]\\
&=\sum_{k=1}^{n}\phi_n(k)\,E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(i)\right)\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p+1}(k)\right)\right]\\
&=\sum_{k=1}^{n}\sum_{l=1}^{n}\phi_n(k)\,\phi_n(l)\,E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(i)\right)\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(l)\right)\right]\\
&=\sum_{l=1}^{n}\phi_n(l)\,E\left[\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(i)\right)\boldsymbol{v}_t\left(\underbrace{\delta_t\cdots\delta_t}_{p}(l)\right)\right]
\end{aligned}
$$

which is the same as for $r=p+1$. It should be clear that this same process would apply for all $r\geq p+2$, which establishes the result.

4 Proof of theorem 1.

The state vector of interest and its law of motion are conjectured to be:

$$X_t \equiv \begin{bmatrix} \boldsymbol{x}_t \\ \overline{E}_t [X_t] \\ \overset{1:\sim}{E}_t [X_t] \\ \overset{2:\sim}{E}_t [X_t] \\ \vdots \end{bmatrix} = F X_{t-1} + G_1 \boldsymbol{u}_t + G_2 \boldsymbol{z}_t + G_3 \boldsymbol{e}_t + G_4 \boldsymbol{z}_{t-1} \tag{10}$$

while agents' private/public and social signals are given by:

$$\boldsymbol{s}_t^p (i) = D_1 \boldsymbol{x}_t + D_2 X_{t-1} + R_1 \boldsymbol{v}_t (i) + R_2 \boldsymbol{e}_t + R_3 \boldsymbol{z}_{t-1} \tag{11a}$$

$$\boldsymbol{s}_t^s (i) = \boldsymbol{\lambda}_1' E_{t-1} (\delta_{t-1} (i)) [X_{t-1}] + \boldsymbol{\lambda}_2' \boldsymbol{x}_{t-1} + \boldsymbol{\lambda}_3' \boldsymbol{v}_{t-1} (\delta_{t-1} (i)) \tag{11b}$$

Together, these describe a linear state space system to which a Kalman filter provides the optimal linear estimator (in the sense of minimising mean squared error).

As discussed in the main text, the system described here is not in the form of a classic state space problem, both because of the presence of the lagged state in agents' signals and because of the moving average component of the law of motion. Lemma 1 demonstrated that we do not need to include $\boldsymbol{z}_t$ in the agents' state vector of interest. To deal with the lagged observations, we follow Nimark (2011b) in developing a modified Kalman filter that does not require the stacking of the state vectors of interest.

To begin, we define the matrices S_x, T_s and T_{w_p} as the matrices that select $\boldsymbol{x}_t$, $\overline{E}_t [X_t]$ and $\overset{p:\sim}{E}_t [X_t]$ respectively from X_t (e.g., $T_{w_2} X_t = \overset{2:\sim}{E}_t [X_t]$).

We also define the general notation that $\theta_{t|q}^{\text{err}} (i)$ represents the error in agent i's period-q expectation regarding θ_t. In particular, we will use the following:

$$\begin{aligned} \boldsymbol{s}_{t|t-1}^{\text{err}} (i) &\equiv \boldsymbol{s}_t (i) - E_{t-1} (i) [\boldsymbol{s}_t (i)] &&: \text{signal innovation} \\ X_{t|t-1}^{\text{err}} (i) &\equiv X_t - E_{t-1} (i) [X_t] &&: \text{prior error} \\ X_{t|t}^{\text{err}} (i) &\equiv X_t - E_t (i) [X_t] &&: \text{contemporaneous error} \end{aligned}$$

4.1 The filter

We proceed by deploying a Gram-Schmidt orthogonalisation of agents' signals. That is, noting that the signal innovation

$$\boldsymbol{s}_{t|t-1}^{\text{err}} (i) \equiv \boldsymbol{s}_t (i) - E_{t-1} (i) [\boldsymbol{s}_t (i)] \tag{12}$$

contains only *new* information available to i in period t, we conclude that it must be orthogonal to any of j's estimates based on information from earlier periods. We can therefore use the standard result that $E[x|y,z] = E[x|y] + E[x|z]$ when $y \perp z$, so that

$$\begin{aligned} E_t (i) [X_t] &= E [X_t | \mathcal{I}_{t-1} (i)] + E \left[X_t | \boldsymbol{s}_{t|t-1}^{\text{err}} (i) \right] \\ &= E_{t-1} (i) [X_t] + K_t \boldsymbol{s}_{t|t-1}^{\text{err}} (i) \end{aligned} \tag{13}$$

for some projection matrix, K_t (the Kalman gain). Note that K_t does not require an agent subscript as the problem is symmetric for all agents.

Optimality then requires that the projection matrix, K_t, be such that the signal innovation, $\boldsymbol{s}^{\text{err}}_{t|t-1}(i)$, is orthogonal to the projection error, $X_t - K_t\boldsymbol{s}^{\text{err}}_{t|t-1}(i)$. That is, we require that

$$E\left[\left(X_t - K_t\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\right)\boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] = 0$$

Rearranging then gives an expression for the optimal Kalman gain:

$$K_t = E\left[X_t\boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right]\left(E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\,\boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right]\right)^{-1}\ \forall i \tag{14}$$

which, since the unconditional expectations of X_t and all signal innovations are zero, is simply

$$K_t = Cov(X_t, \boldsymbol{s}^{\text{err}}_{t|t-1}(i))\left[Var\left(\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\right)\right]^{-1}$$

In order to evaluate this, it is necessary to construct expressions for the innovation in agents' private and social signals. We consider each in turn.

Agents' private signals

To begin, we substitute the conjectured state law of motion into the private signal equation to get:

$$\begin{aligned}\boldsymbol{s}^p_t(j) &= (D_1S_xF + D_2)X_{t-1} + D_1S_xG_1\boldsymbol{u}_t \\ &\quad + R_1\boldsymbol{v}_t(j) + R_2\boldsymbol{e}_t + R_3\boldsymbol{z}_{t-1}\end{aligned} \tag{15}$$

where we have used the fact that $\boldsymbol{x}_t$ is independent of network shocks to ignore the $G_2\boldsymbol{z}_t$ and $G_4\boldsymbol{z}_{t-1}$ components of X_t. From this, we see that i's prior expectation of her private signal will be given by

$$E_{t-1}(i)\left[\boldsymbol{s}^p_t(i)\right] = (D_1S_xF + D_2)E_{t-1}(i)\left[X_{t-1}\right] \tag{16}$$

where we have made use of lemma 1 to drop the term in $E_{t-1}(i)\left[\boldsymbol{z}_{t-1}\right]$. Subtracting equation (16) from (15) then gives the innovation in agents' private signals as

$$\begin{aligned}\boldsymbol{s}^p_{t|t-1}(i) &= (D_1S_xF + D_2)X^{\text{err}}_{t-1|t-1}(i) + D_1S_xG_1\boldsymbol{u}_t \\ &\quad + R_1\boldsymbol{v}_t(j) + R_2\boldsymbol{e}_t + R_3\boldsymbol{z}_{t-1}\end{aligned} \tag{17}$$

where $X^{\text{err}}_{t|t}(i)$ is i's contemporaneous error in estimating X_t.

Agents' social signals

For the social signal, and assuming temporarily that agents observe the actions of only one competitor, we make use of proposition 1 to write the prior expectation as

$$E_{t-1}(i)\left[\boldsymbol{s}^s_t(i)\right] = \boldsymbol{\lambda}_1'E_{t-1}(i)\left[\widetilde{E}_{t-1}\left[X_{t-1}\right]\right] + \boldsymbol{\lambda}_2'E_{t-1}(i)\left[\boldsymbol{x}_{t-1}\right] + \boldsymbol{\lambda}_3'E_{t-1}(i)\left[\widetilde{\boldsymbol{v}}_{t-1}\right]$$

Given that $E_t(i)\left[\boldsymbol{z}_t\right] = 0$, $S_xX_t = \boldsymbol{x}_t$ and $T_{w_1}X_t = {}^{\{1\}}\widetilde{E}_t\left[X_t\right]$, we can write this as

$$E_{t-1}(i)\left[\boldsymbol{s}^s_t(i)\right] = \left(\boldsymbol{\lambda}_2'S_x + \boldsymbol{\lambda}_1'T_w\right)E_{t-1}(i)\left[X_{t-1}\right] \tag{18}$$

Subtracting (18) from (11b), we then have that the innovation in the agent's social signal is given by:

$$\begin{aligned}\boldsymbol{s}^s_{t|t-1}(i) &= \boldsymbol{\lambda}_2'S_xX^{\text{err}}_{t-1|t-1}(i) \\ &\quad + \boldsymbol{\lambda}_1'E_{t-1}(\delta_{t-1}(i))\left[X_{t-1}\right] - \boldsymbol{\lambda}_1'T_wE_{t-1}(i)\left[X_{t-1}\right] \\ &\quad + \boldsymbol{\lambda}_3'\boldsymbol{v}_{t-1}(\delta_{t-1}(i))\end{aligned}$$

Adding and subtracting $\boldsymbol{\lambda}_1' T_w X_{t-1}$ on the right-hand side then gives

$$\begin{aligned} \boldsymbol{s}^s_{t|t-1}(i) = &\left(\boldsymbol{\lambda}_2' S_x + \boldsymbol{\lambda}_1' T_w\right) X^{\text{err}}_{t-1|t-1}(i) \\ &- \boldsymbol{\lambda}_1' \left(T_w X_{t-1} - E_{t-1}\left(\delta_{t-1}(i)\right)[X_{t-1}]\right) \\ &+ \boldsymbol{\lambda}_3' \boldsymbol{v}_{t-1}\left(\delta_{t-1}(i)\right) \end{aligned}$$

and finally now adding and subtracting $\boldsymbol{\lambda}_1' X_{t-1}$ on the right-hand side gives

$$\begin{aligned} \boldsymbol{s}^s_{t|t-1}(i) = &\left(\boldsymbol{\lambda}_2' S_x + \boldsymbol{\lambda}_1' T_w\right) X^{\text{err}}_{t-1|t-1}(i) - \boldsymbol{\lambda}_1' X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i)\right) \\ &+ \boldsymbol{\lambda}_1' \left(I - T_w\right) X_{t-1} \\ &+ \boldsymbol{\lambda}_3' \boldsymbol{v}_{t-1}\left(\delta_{t-1}(i)\right) \end{aligned}$$

Crucially, we have that the innovation in i's social signal includes not only a term in their own contemporaneous error from the previous period but also a term in their *observee*'s error.

The combined signal innovation

Stacking the private, public and social signal innovations, we then obtain

$$\begin{aligned} \boldsymbol{s}^{\text{err}}_{t|t-1}(i) = &M_1 X^{\text{err}}_{t-1|t-1}(i) + M_2 X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i)\right) + M_3 X_{t-1} \\ &+ N_1 \boldsymbol{u}_t + N_2 \boldsymbol{v}_t(i) + N_3 \boldsymbol{e}_t + N_4 \boldsymbol{v}_{t-1}\left(\delta_{t-1}(i)\right) + N_5 \boldsymbol{z}_{t-1} \end{aligned} \tag{19a}$$

where

$$M_1 = \begin{bmatrix} D_1 S_x F + D_2 \\ \boldsymbol{\lambda}_2' S_x + \boldsymbol{\lambda}_1' T_w \end{bmatrix} \quad M_2 = \begin{bmatrix} \mathbf{0} \\ -\boldsymbol{\lambda}_1' \end{bmatrix} \quad M_3 = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\lambda}_1' \left(I - T_w\right) \end{bmatrix} \tag{19b}$$

$$N_1 = \begin{bmatrix} D_1 S_x G_1 \\ \mathbf{0} \end{bmatrix} \quad N_2 = \begin{bmatrix} R_1 \\ \mathbf{0} \end{bmatrix} \quad N_3 = \begin{bmatrix} R_2 \\ \mathbf{0} \end{bmatrix} \quad N_4 = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\lambda}_3' \end{bmatrix} \quad N_5 = \begin{bmatrix} R_3 \\ \mathbf{0} \end{bmatrix} \tag{19c}$$

Considering two or more observees is then obtained by further stacking the signals

$$\begin{aligned} \boldsymbol{s}^{\text{err}}_{t|t-1}(i) = &M_1 X^{\text{err}}_{t-1|t-1}(i) + M_2 \begin{bmatrix} X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i,1)\right) \\ X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i,2)\right) \end{bmatrix} + M_3 X_{t-1} \\ &+ N_1 \boldsymbol{u}_t + N_2 \boldsymbol{v}_t(i) + N_3 \boldsymbol{e}_t + N_4 \begin{bmatrix} \boldsymbol{v}_{t-1}\left(\delta_{t-1}(i,1)\right) \\ \boldsymbol{v}_{t-1}\left(\delta_{t-1}(i,2)\right) \end{bmatrix} + N_5 \boldsymbol{z}_{t-1} \end{aligned} \tag{20a}$$

where

$$M_1 = \begin{bmatrix} D_1 S_x F + D_2 \\ \boldsymbol{\lambda}_2' S_x + \boldsymbol{\lambda}_1' T_w \\ \boldsymbol{\lambda}_2' S_x + \boldsymbol{\lambda}_1' T_w \end{bmatrix} \quad M_2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ -\boldsymbol{\lambda}_1' & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{\lambda}_1' \end{bmatrix} \quad M_3 = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\lambda}_1' \left(I - T_w\right) \\ \boldsymbol{\lambda}_1' \left(I - T_w\right) \end{bmatrix} \tag{20b}$$

$$N_1 = \begin{bmatrix} D_1 S_x G_1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad N_2 = \begin{bmatrix} R_1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad N_3 = \begin{bmatrix} R_2 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad N_4 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \boldsymbol{\lambda}_3' & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\lambda}_3' \end{bmatrix} \quad N_5 = \begin{bmatrix} R_3 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{20c}$$

For the remainder of this appendix, we shall use the notation of a single observee on the understanding that the signal innovation may be replace as above for an arbitrary number of competitors observed.

Deriving the Kalman gain

We first expand the first term in equation (14) as

$$
E\left[X_t \boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] = E\left[\begin{array}{c} (FX_{t-1} + G_1\boldsymbol{u}_t + G_2\boldsymbol{z}_t + G_4\boldsymbol{z}_{t-1} + G_3\boldsymbol{e}_t) \\ \times \left(\begin{array}{l} M_1 X^{\text{err}}_{t-1|t-1}(i) \\ +M_2 X^{\text{err}}_{t-1|t-1}(\delta_{t-1}(i)) \\ +M_3 X_{t-1} \\ +N_1\boldsymbol{u}_t + N_2\boldsymbol{v}_t(i) + N_3\boldsymbol{e}_t \\ +N_4\boldsymbol{v}_{t-1}(\delta_{t-1}(i)) + N_5\boldsymbol{z}_{t-1} \end{array}\right)' \end{array}\right]
$$

$$
= E\left[\begin{array}{l} (FX_{t-1})\left(M_1 X^{\text{err}}_{t-1|t-1}(i)\right)' \\ +(FX_{t-1})\left(M_2 X^{\text{err}}_{t-1|t-1}(\delta_{t-1}(i))\right)' \\ +(FX_{t-1})(M_3 X_{t-1})' \\ +(FX_{t-1})(N_5\boldsymbol{z}_{t-1})' \\ +(G_1\boldsymbol{u}_t)(N_1\boldsymbol{u}_t)' \\ +(G_3\boldsymbol{e}_t)(N_3\boldsymbol{e}_t)' \\ +(G_4\boldsymbol{z}_{t-1})\left(M_1 X^{\text{err}}_{t-1|t-1}(i)\right)' \\ +(G_4\boldsymbol{z}_{t-1})\left(M_2 X^{\text{err}}_{t-1|t-1}(\delta_{t-1}(i))\right)' \\ +(G_4\boldsymbol{z}_{t-1})(M_3 X_{t-1})' \\ +(G_4\boldsymbol{z}_{t-1})(N_5\boldsymbol{z}_{t-1})' \end{array}\right] \tag{21}
$$

where we use the fact that period-t shocks are orthogonal to period-$(t-1)$ objects and make use of assumption 2 (which grants us that $\lim_{n\to\infty}\phi_n(i) = 0\ \forall i$) to note that there is no covariance between period-$(t-1)$ objects and $\boldsymbol{v}_{t-1}(i)\ \forall i$.

next, we note that for any j and any t, we may write

$$
\begin{aligned}
E\left[X_t X^{\text{err}}_{t|t}(j)'\right] &= E\left[\left(X^{\text{err}}_{t|t}(j) + E_t(j)[X_t]\right) X^{\text{err}}_{t|t}(j)'\right] \\
&= E\left[X^{\text{err}}_{t|t}(j)\, X^{\text{err}}_{t|t}(j)'\right] \\
&= V_{t|t}
\end{aligned}
$$

where the second equality makes use of the fact that since $E_t(j)[X_t]$ is spanned by the set of orthogonal signal innovations $\left\{\boldsymbol{s}^{\text{err}}_{t|t-1}(j), \boldsymbol{s}^{\text{err}}_{t-1|t-2}(j), \cdots\right\}$ and these are orthogonal to $X^{\text{err}}_{t|t}(j)$ by construction, then it must be that $E_t(j)[X_t]$ and $X^{\text{err}}_{t|t}(j)$ are orthogonal for all j and t. Note that $V_{t|t} \equiv E\left[X^{\text{err}}_{t|t}(j)\, X^{\text{err}}_{t|t}(j)'\right]\ \forall j$ is the variance of each agent's contemporaneous error (common to all agents as their problems are symmetric).

Using this, we may rewrite (21) as

$$
\begin{aligned}
E\left[X_t \boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] &= FV_{t-1|t-1}M_1' \\
&+ FV_{t-1|t-1}M_2' \\
&+ FU_{t-1}M_3' \\
&+ FG_2\Sigma_{zz}N_5' \\
&+ G_1\Sigma_{uu}N_1' \\
&+ G_3\Sigma_{ee}N_3' \\
&+ G_4\Sigma_{zz}G_2'(M_1 + M_2 + M_3)' \\
&+ G_4\Sigma_{zz}N_5'
\end{aligned}
$$

or, defining $M \equiv \begin{bmatrix} M_1 & M_2 & M_3 \end{bmatrix}$, as simply

$$
\begin{aligned}
E\left[X_t \boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] = F &\begin{bmatrix} V_{t-1|t-1} & V_{t-1|t-1} & U_{t-1} \end{bmatrix} M' \\
&+ G_1 \Sigma_{uu} N_1' \\
&+ F G_2 \Sigma_{zz} N_5' \\
&+ G_3 \Sigma_{ee} N_3' \\
&+ G_4 \Sigma_{zz} G_2' \left(M_1 + M_2 + M_3\right)' \\
&+ G_4 \Sigma_{zz} N_5'
\end{aligned} \tag{22}
$$

Turning to the second term in equation (14), we have that

$$
\begin{aligned}
E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\, \boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] &= E \left[\begin{array}{l} \begin{pmatrix} \quad M_1 X^{\text{err}}_{t-1|t-1}(i) \\ +M_2 X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i)\right) \\ +M_3 X_{t-1} \\ +N_1 \boldsymbol{u}_t + N_2 \boldsymbol{v}_t(i) \\ +N_4 \boldsymbol{v}_{t-1}\left(\delta_{t-1}(i)\right) + N_5 \boldsymbol{z}_{t-1} + N_3 \boldsymbol{e}_t \end{pmatrix} \\ \times \begin{pmatrix} \quad M_1 X^{\text{err}}_{t-1|t-1}(i) \\ +M_2 X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i)\right) \\ +M_3 X_{t-1} \\ +N_1 \boldsymbol{u}_t + N_2 \boldsymbol{v}_t(i) \\ +N_4 \boldsymbol{v}_{t-1}\left(\delta_{t-1}(i)\right) + N_5 \boldsymbol{z}_{t-1} + N_3 \boldsymbol{e}_t \end{pmatrix}' \end{array} \right] \\
&= E \left[\begin{array}{l} \begin{pmatrix} \quad M_1 X^{\text{err}}_{t-1|t-1}(i) \\ +M_2 X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i)\right) \\ +M_3 X_{t-1} \\ +N_5 \boldsymbol{z}_{t-1} \end{pmatrix} \\ \times \begin{pmatrix} \quad M_1 X^{\text{err}}_{t-1|t-1}(i) \\ +M_2 X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i)\right) \\ +M_3 X_{t-1} \\ +N_5 \boldsymbol{z}_{t-1} \end{pmatrix}' \end{array} \right] \\
&+ M_2 E\left[X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i)\right) \boldsymbol{v}_{t-1}\left(\delta_{t-1}(i)\right)'\right] N_4' \\
&+ N_4 E\left[\boldsymbol{v}_{t-1}\left(\delta_{t-1}(i)\right) X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}(i)\right)'\right] M_2' \\
&+ N_1 \Sigma_{uu} N_1' + N_2 \Sigma_{vv} N_2' + N_4 \Sigma_{vv} N_4' + N_3 \Sigma_{ee} N_3'
\end{aligned}
$$

Expanding out the various cross-products then gives us

$$
\begin{aligned}
E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\,\boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] = {} & M_1 V_{t-1|t-1} M_1' + M_1 W_{t-1|t-1} M_2' + M_1 V_{t-1|t-1} M_3' \\
& + M_2 W_{t-1|t-1} M_1' + M_2 V_{t-1|t-1} M_2' + M_2 V_{t-1|t-1} M_3' \\
& + M_3 V_{t-1|t-1} M_1' + M_3 V_{t-1|t-1} M_2' + M_3 U_{t-1} M_3' \\
& - M_2 K_{t-1} N_2 \Sigma_{vv} N_4' \\
& - N_4 \Sigma_{vv} N_2' K_{t-1}' M_2' \\
& + N_1 \Sigma_{uu} N_1' + N_2 \Sigma_{vv} N_2' + N_4 \Sigma_{vv} N_4' \\
& + (M_1 + M_2 + M_3)\, G_2 \Sigma_{zz} N_5' \\
& + N_5 \Sigma_{zz} G_2' (M_1 + M_2 + M_3)' \\
& + N_3 \Sigma_{ee} N_3'
\end{aligned}
$$

where $W_{t|t} \equiv E\left[X^{\text{err}}_{t|t}(i)\, X^{\text{err}}_{t|t}(j)'\right]\ \forall i \neq j$ is the covariance between any two agents' contemporaneous errors (common to all agent-pairs as their problems are symmetric and the network is opaque so they each have the same probability of observing the same target). Similarly to the covariance term, this may be written simply as

$$
\begin{aligned}
E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\,\boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] = {} & M \begin{bmatrix} V_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \\ W_{t-1|t-1} & V_{t-1|t-1} & V_{t-1|t-1} \\ V_{t-1|t-1} & V_{t-1|t-1} & U_{t-1} \end{bmatrix} M' \\
& - M_2 K_{t-1} N_2 \Sigma_{vv} N_4' \\
& - N_4 \Sigma_{vv} N_2' K_{t-1}' M_2' \\
& + N_1 \Sigma_{uu} N_1' + N_2 \Sigma_{vv} N_2' + N_4 \Sigma_{vv} N_4' \\
& + (M_1 + M_2 + M_3)\, G_2 \Sigma_{zz} N_5' \\
& + N_5 \Sigma_{zz} G_2' (M_1 + M_2 + M_3)' \\
& + N_3 \Sigma_{ee} N_3' \qquad (23)
\end{aligned}
$$

Substituting (22) and (23) into (14) and gathering like terms, we arrive at:

$$
K_t = \begin{pmatrix}
F \begin{bmatrix} V_{t-1|t-1} & V_{t-1|t-1} & U_{t-1} \end{bmatrix} M' \\
+G_1 \Sigma_{uu} N_1' \\
+F G_2 \Sigma_{zz} N_5' \\
+G_4 \Sigma_{zz} G_2' (M_1 + M_2 + M_3)' \\
+G_4 \Sigma_{zz} N_5' \\
+G_3 \Sigma_{ee} N_3'
\end{pmatrix}
\times \begin{bmatrix}
M \begin{bmatrix} V_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \\ W_{t-1|t-1} & V_{t-1|t-1} & V_{t-1|t-1} \\ V_{t-1|t-1} & V_{t-1|t-1} & U_{t-1} \end{bmatrix} M' \\
+(M_1 + M_2 + M_3)\, G_2 \Sigma_{zz} N_5' \\
+N_5 \Sigma_{zz} G_2' (M_1 + M_2 + M_3)' \\
-M_2 K_{t-1} N_2 \Sigma_{vv} N_4' \\
-N_4 \Sigma_{vv} N_2' K_{t-1}' M_2' \\
+N_1 \Sigma_{uu} N_1' + N_2 \Sigma_{vv} N_2' + N_4 \Sigma_{vv} N_4' + N_3 \Sigma_{ee} N_3'
\end{bmatrix}^{-1} \qquad (24)
$$

4.2 Evolution of the variance-covariance matricies

Unconditional variance of the state vector of interest

From the conjectured law of motion, we can read immediately that the variance of the state vector of interest evolves as:

$$\begin{aligned} U_t &= F U_{t-1} F' \\ &\quad + G_1 \Sigma_{uu} G_1' + G_2 \Sigma_{zz} G_2' + G_3 \Sigma_{ee} G_3' + G_4 \Sigma_{zz} G_4' + F G_2 \Sigma_{zz} G_4' + G_4 \Sigma_{zz} G_2' F' \end{aligned} \tag{25}$$

Variance of agents' expectation errors

First, subtracting $E_{t-1}(i)[X_t]$ from each side of the state equation, we have:

$$\begin{aligned} X_t - E_{t-1}(i)[X_t] &= F\left(X_{t-1} - E_{t-1}(i)[X_{t-1}]\right) \\ &\quad + G_1 \boldsymbol{u}_t + G_2 \boldsymbol{z}_t + G_3 \boldsymbol{e}_t + G_4 \boldsymbol{z}_{t-1} \end{aligned} \tag{26}$$

Taking the variance of each side, we have that the prior variance will be given by:

$$\begin{aligned} V_{t|t-1} &= F V_{t-1|t-1} F' \\ &\quad + G_1 \Sigma_{uu} G_1' + G_2 \Sigma_{zz} G_2' + G_3 \Sigma_{ee} G_3' + G_4 \Sigma_{zz} G_4' + F G_2 \Sigma_{zz} G_4' + G_4 \Sigma_{zz} G_2' F' \end{aligned} \tag{27}$$

next, we subtract each side of equation (13) from X_t and rearrange to obtain

$$\left(X_t - E_t(i)[X_t]\right) + K_t \boldsymbol{s}^{\text{err}}_{t|t-1}(i) = \left(X_t - E_{t-1}(i)[X_t]\right) \tag{28}$$

Since the signal innovation is orthogonal to the contemporaneous error, $X_t - E_t(i)[X_t]$ by construction, the variance of the right-hand side must equal the sum of the variances on the left-hand side, thereby giving:

$$V_{t|t} + K_t Var\left(\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\right) K_t' = V_{t|t-1}$$

or

$$V_{t|t} = V_{t|t-1} - K_t \left(\begin{array}{l} M \begin{bmatrix} V_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \\ W_{t-1|t-1} & V_{t-1|t-1} & V_{t-1|t-1} \\ V_{t-1|t-1} & V_{t-1|t-1} & U_{t-1} \end{bmatrix} M' \\ + (M_1 + M_2 + M_3) G_2 \Sigma_{zz} N_5' \\ + N_5 \Sigma_{zz} G_2' (M_1 + M_2 + M_3)' \\ - M_2 K_{t-1} N_2 \Sigma_{vv} N_4' \\ - N_4 \Sigma_{vv} N_2' K_{t-1}' M_2' \\ + N_1 \Sigma_{uu} N_1' + N_2 \Sigma_{vv} N_2' + N_3 \Sigma_{ee} N_3' + N_4 \Sigma_{vv} N_4' \end{array} \right) K_t' \tag{29}$$

Covariance between agents' expectation errors

First, from (26), we have that the prior covariance between two agents' errors is given by:

$$\begin{aligned} W_{t|t-1} &\equiv E\left[X^{\text{err}}_{t|t-1}(i)\, X_{t|t-1}(j)'\right] \\ &= F W_{t-1|t-1} F' \\ &\quad + G_1 \Sigma_{uu} G_1' + G_2 \Sigma_{zz} G_2' + G_3 \Sigma_{ee} G_3' + G_4 \Sigma_{zz} G_4' + F G_2 \Sigma_{zz} G_4' + G_4 \Sigma_{zz} G_2' F' \end{aligned} \tag{30}$$

next, returning to equation (28)

$$\left(X_t - E_t\left(i\right)\left[X_t\right]\right) = \left(X_t - E_{t-1}\left(i\right)\left[X_t\right]\right) - K_t \boldsymbol{s}^{\text{err}}_{t|t-1}\left(i\right) \tag{31}$$

note that agent i's signal innovation will not necessarily be orthogonal to either of j's expectation errors, so we expand this fully to obtain

$$\begin{aligned} W_{t|t} = {} & W_{t|t-1} \\ & + K_t\, Cov\left(\boldsymbol{s}^{\text{err}}_{t|t-1}\left(i\right), \boldsymbol{s}^{\text{err}}_{t|t-1}\left(j\right)\right) K_t' \\ & - Cov\left(X^{\text{err}}_{t|t-1}\left(i\right), \boldsymbol{s}^{\text{err}}_{t|t-1}\left(j\right)\right) K_t' \\ & - K_t\, Cov\left(\boldsymbol{s}^{\text{err}}_{t|t-1}\left(i\right), X_{t|t-1}\left(j\right)\right) \end{aligned} \tag{32}$$

For the second term on the right-hand side, we have

$$\begin{aligned} E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}\left(i\right)\boldsymbol{s}^{\text{err}}_{t|t-1}\left(j\right)'\right] &= E\left[\begin{array}{l} \begin{pmatrix} M_1 X^{\text{err}}_{t-1|t-1}\left(i\right) \\ +M_2 X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}\left(i\right)\right) \\ +M_3 X_{t-1} \\ +N_1 \boldsymbol{u}_t + N_2 \boldsymbol{v}_t\left(i\right) \\ +N_4 \boldsymbol{v}_{t-1}\left(\delta_{t-1}\left(i\right)\right) + N_5 \boldsymbol{z}_{t-1} + N_3 \boldsymbol{e}_t \end{pmatrix} \\ \times \begin{pmatrix} M_1 X_{t-1|t-1}\left(j\right) \\ +M_2 X_{t-1|t-1}\left(\delta_{t-1}\left(j\right)\right) \\ +M_3 X_{t-1} \\ +N_1 \boldsymbol{u}_t + N_2 \boldsymbol{v}_t\left(j\right) \\ +N_4 \boldsymbol{v}_{t-1}\left(\delta_{t-1}\left(j\right)\right) + N_5 \boldsymbol{z}_{t-1} + N_3 \boldsymbol{e}_t \end{pmatrix}' \end{array}\right] \\ &= E\left[\begin{array}{l} \begin{pmatrix} M_1 X^{\text{err}}_{t-1|t-1}\left(i\right) \\ +M_2 X^{\text{err}}_{t-1|t-1}\left(\delta_{t-1}\left(i\right)\right) \\ +M_3 X_{t-1} \\ +N_5 \boldsymbol{z}_{t-1} \end{pmatrix} \\ \times \begin{pmatrix} M_1 X_{t-1|t-1}\left(j\right) \\ +M_2 X_{t-1|t-1}\left(\delta_{t-1}\left(j\right)\right) \\ +M_3 X_{t-1} \\ +N_5 \boldsymbol{z}_{t-1} \end{pmatrix}' \end{array}\right] \\ &\quad + N_1 \Sigma_{uu} N_1' \\ &\quad + N_3 \Sigma_{ee} N_3' \end{aligned}$$

Given $i \neq j$ and assumption 2, it must be the case that i, j, $\delta_{t-1}\left(i\right)$ and $\delta_{t-1}\left(j\right)$ are four different agents, almost surely. We therefore have

$$\begin{aligned} E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}\left(i\right)\boldsymbol{s}^{\text{err}}_{t|t-1}\left(j\right)'\right] &= M \begin{bmatrix} W_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \\ W_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \\ V_{t-1|t-1} & V_{t-1|t-1} & U_{t-1} \end{bmatrix} M' \\ &\quad + \left(M_1 + M_2 + M_3\right) G_2 \Sigma_{zz} N_5' \\ &\quad + N_5 \Sigma_{zz} G_2' \left(M_1 + M_2 + M_3\right)' \\ &\quad + N_1 \Sigma_{uu} N_1' \\ &\quad + N_3 \Sigma_{ee} N_3' \end{aligned} \tag{33}$$

For the third term, we have

$$Cov\left(X^{\text{err}}_{t|t-1}(i), \boldsymbol{s}^{\text{err}}_{t|t-1}(j)\right) = E\left[\begin{array}{c}\left(\begin{array}{l} FX_{t-1|t-1}(j) \\ +G_1\boldsymbol{u}_t \\ +G_2\boldsymbol{z}_t \\ +G_4\boldsymbol{z}_{t-1} \\ +G_3\boldsymbol{e}_t \end{array}\right) \\ \times\left(\begin{array}{l} M_1X^{\text{err}}_{t-1|t-1}(i) \\ +M_2X^{\text{err}}_{t-1|t-1}(\delta_{t-1}(i)) \\ +M_3X_{t-1} \\ +N_1\boldsymbol{u}_t + N_2\boldsymbol{v}_t(i) \\ +N_4\boldsymbol{v}_{t-1}(\delta_{t-1}(i)) + N_5\boldsymbol{z}_{t-1} + N_3\boldsymbol{e}_t \end{array}\right)' \end{array}\right]$$

$$\begin{aligned} &= F\begin{bmatrix} V_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \end{bmatrix} M' \\ &+ G_1\Sigma_{uu}N_1' \\ &+ FG_2\Sigma_{zz}N_5' \\ &+ G_4\Sigma_{zz}G_2'(M_1+M_2+M_3)' \\ &+ G_4\Sigma_{zz}N_5' \\ &+ G_3\Sigma_{zz}N_3' \end{aligned} \tag{34}$$

while the fourth term is the transpose of the same.

Filter summary

In summary, the filter evolves through the following system of equations:

$$\begin{aligned} E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\,\boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] = M &\begin{bmatrix} V_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \\ W_{t-1|t-1} & V_{t-1|t-1} & V_{t-1|t-1} \\ V_{t-1|t-1} & V_{t-1|t-1} & U_{t-1} \end{bmatrix} M' \\ &+ (M_1+M_2+M_3)\,G_2\Sigma_{zz}N_5' \\ &+ N_5\Sigma_{zz}G_2'(M_1+M_2+M_3)' \\ &- M_2K_{t-1}N_2\Sigma_{vv}N_4' \\ &- N_4\Sigma_{vv}N_2'K_{t-1}'M_2' \\ &+ N_1\Sigma_{uu}N_1' + N_2\Sigma_{vv}N_2' + N_4\Sigma_{vv}N_4' \end{aligned} \tag{35a}$$

$$\begin{aligned} E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\,\boldsymbol{s}^{\text{err}}_{t|t-1}(j)'\right] = M &\begin{bmatrix} W_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \\ W_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \\ V_{t-1|t-1} & V_{t-1|t-1} & U_{t-1} \end{bmatrix} M' \\ &+ (M_1+M_2+M_3)\,G_2\Sigma_{zz}N_5' \\ &+ N_5\Sigma_{zz}G_2'(M_1+M_2+M_3)' \\ &+ N_1\Sigma_{uu}N_1' \end{aligned} \tag{35b}$$

$$
\begin{aligned}
E\left[X_t \boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] = F &\begin{bmatrix} V_{t-1|t-1} & V_{t-1|t-1} & U_{t-1} \end{bmatrix} M' \\
&+ G_1 \Sigma_{uu} N_1' \\
&+ F G_2 \Sigma_{zz} N_5' \\
&+ G_4 \Sigma_{zz} G_2' (M_1 + M_2 + M_3)' \\
&+ G_4 \Sigma_{zz} N_5'
\end{aligned} \tag{35c}
$$

$$
\begin{aligned}
E\left[X^{\text{err}}_{t|t-1}(i)\, \boldsymbol{s}^{\text{err}}_{t|t-1}(j)'\right] = F &\begin{bmatrix} V_{t-1|t-1} & W_{t-1|t-1} & V_{t-1|t-1} \end{bmatrix} M' \\
&+ G_1 \Sigma_{uu} N_1' \\
&+ F G_2 \Sigma_{zz} N_5' \\
&+ G_4 \Sigma_{zz} G_2' (M_1 + M_2 + M_3)' \\
&+ G_4 \Sigma_{zz} N_5'
\end{aligned} \tag{35d}
$$

$$
K_t = E\left[X_t \boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] \left(E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\, \boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] \right)^{-1} \tag{35e}
$$

$$
\begin{aligned}
U_t &= F U_{t-1} F' \\
&\quad + G_1 \Sigma_{uu} G_1' + G_2 \Sigma_{zz} G_2' + G_4 \Sigma_{zz} G_4' + F G_2 \Sigma_{zz} G_4' + G_4 \Sigma_{zz} G_2' F'
\end{aligned} \tag{35f}
$$

$$
\begin{aligned}
V_{t|t-1} &= F V_{t-1|t-1} F' \\
&\quad + G_1 \Sigma_{uu} G_1' + G_2 \Sigma_{zz} G_2' + G_4 \Sigma_{zz} G_4' + F G_2 \Sigma_{zz} G_4' + G_4 \Sigma_{zz} G_2' F'
\end{aligned} \tag{35g}
$$

$$
\begin{aligned}
W_{t|t-1} &= F W_{t-1|t-1} F' \\
&\quad + G_1 \Sigma_{uu} G_1' + G_2 \Sigma_{zz} G_2' + G_4 \Sigma_{zz} G_4' + F G_2 \Sigma_{zz} G_4' + G_4 \Sigma_{zz} G_2' F'
\end{aligned} \tag{35h}
$$

$$
V_{t|t} = V_{t|t-1} - K_t\, E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\, \boldsymbol{s}^{\text{err}}_{t|t-1}(i)'\right] K_t' \tag{35i}
$$

$$
\begin{aligned}
W_{t|t} = W_{t|t-1} &+ K_t E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\, \boldsymbol{s}^{\text{err}}_{t|t-1}(j)'\right] K_t' \\
&- E\left[X^{\text{err}}_{t|t-1}(i)\, \boldsymbol{s}^{\text{err}}_{t|t-1}(j)'\right] K_t' \\
&- K_t E\left[\boldsymbol{s}^{\text{err}}_{t|t-1}(i)\, X_{t|t-1}(j)'\right]
\end{aligned} \tag{35j}
$$

Provided that all eigenvalues of F are within the unit circle, then there will exist a steady state (i.e. time-invariant) filter, found by iterating these equations forward until convergence is achieved.

4.3 Confirming the conjectured law of motion

The state vector of interest and its law of motion are conjectured to be:

$$X_t \equiv \begin{bmatrix} \boldsymbol{x}_t \\ \overline{E}_t [X_t] \\ \overset{1:\sim}{E}_t [X_t] \\ \overset{2:\sim}{E}_t [X_t] \\ \vdots \end{bmatrix} = F X_{t-1} + G_1 \boldsymbol{u}_t + G_2 \boldsymbol{z}_t + G_3 \boldsymbol{e}_t + G_4 \boldsymbol{z}_{t-1} \tag{36}$$

To confirm this law of motion, we first combining equations (13) and (20) to write the agents' filter as:

$$\begin{aligned} E_t (i) [X_t] = {} & F E_{t-1} (i) [X_{t-1}] \\ & + K \begin{pmatrix} M_1 (X_{t-1} - E_{t-1} (i) [X_{t-1}]) \\ +M_2 (X_{t-1} - E_{t-1} (\delta_{t-1} (i)) [X_{t-1}]) \\ +M_3 X_{t-1} \\ +N_1 \boldsymbol{u}_t + N_2 \boldsymbol{v}_t (i) + N_3 \boldsymbol{e}_t \\ +N_4 \boldsymbol{v}_{t-1} (\delta_{t-1} (i)) + N_5 \boldsymbol{z}_{t-1} \end{pmatrix} \end{aligned}$$

Gathering like terms gives

$$\begin{aligned} E_t (i) [X_t] = {} & K (M_1 + M_2 + M_3) X_{t-1} \\ & + (F - K M_1) E_{t-1} (i) [X_{t-1}] \\ & - K M_2 E_{t-1} (\delta_{t-1} (i)) [X_{t-1}] \\ & + K N_1 \boldsymbol{u}_t \\ & + K N_2 \boldsymbol{v}_t (i) \\ & + K N_3 \boldsymbol{e}_t \\ & + K N_4 \boldsymbol{v}_{t-1} (\delta_{t-1} (i)) \\ & + K N_5 \boldsymbol{z}_{t-1} \end{aligned} \tag{37}$$

Taking the simple average of equation (37) gives

$$\begin{aligned} \overline{E}_t [X_t] = {} & K (M_1 + M_2 + M_3) X_{t-1} \\ & + (F - K M_1) \overline{E}_{t-1} [X_{t-1}] \\ & - K M_2 \overset{1:\sim}{E}_{t-1} [X_{t-1}] \\ & + K N_1 \boldsymbol{u}_t \\ & + K N_3 \boldsymbol{e}_t \\ & + K N_4 \overset{1:\sim}{\tilde{\boldsymbol{v}}}_{t-1} \\ & + K N_5 \boldsymbol{z}_{t-1} \end{aligned}$$

where I have used proposition 1 to replace $\int_0^1 \boldsymbol{v}_{t-1} (\delta_{t-1} (i)) \, di$ with $\overset{1:\sim}{\tilde{\boldsymbol{v}}}_{t-1}$. But since $\overset{1:\sim}{\tilde{\boldsymbol{v}}}_{t-1}$ is part of $\boldsymbol{z}_{t-1}$, while $\overline{E}_{t-1} [X_{t-1}]$ and $\overset{1:\sim}{E}_{t-1} [X_{t-1}]$ are part of X_{t-1}, we can simplify this down to:

$$\begin{aligned} \overline{E}_t [X_t] = {} & \{ K (M_1 + M_2 + M_3) + (F - K M_1) T_s - K M_2 T_{w_1} \} X_{t-1} \\ & + K N_1 \boldsymbol{u}_t \\ & + K N_3 \boldsymbol{e}_t \\ & + K \left(\begin{bmatrix} N_4 & \boldsymbol{0}_{1 \times \infty} \end{bmatrix} + N_5 \right) \boldsymbol{z}_{t-1} \end{aligned} \tag{38}$$

next, taking the p-th weighted average of equation (37) gives

$$\begin{aligned}
\overset{1:\sim}{E}_t [X_t] = &K (M_1 + M_2 + M_3) X_{t-1} \\
&+ (F - KM_1) \overset{1:\sim}{E}_{t-1} [X_{t-1}] \\
&- KM_2 \overset{p+1:\sim}{E}_{t-1} [X_{t-1}] \\
&+ KN_1 \boldsymbol{u}_t \\
&+ KN_2 \overset{p:\sim}{\boldsymbol{v}}_t \\
&+ KN_3 \boldsymbol{e}_t \\
&+ KN_4 \overset{p+1:\sim}{\boldsymbol{v}}_{t-1} \\
&+ KN_5 \boldsymbol{z}_{t-1}
\end{aligned}$$

where the last two terms have again made use of proposition 1. From this, we read immediately that

$$\begin{aligned}
\overset{p:\sim}{E}_t [X_t] = &\{K (M_1 + M_2 + M_3) + (F - KM_1) T_{w_p} - KM_2 T_{w_{p+1}}\} X_{t-1} \\
&+ KN_1 \boldsymbol{u}_t \\
&+ K \begin{bmatrix} \mathbf{0}_{1\times r(q-1)} & N_2 & \mathbf{0}_{1\times\infty} \end{bmatrix} \boldsymbol{z}_t \\
&+ KN_3 \boldsymbol{e}_t \\
&+ K \left(\begin{bmatrix} \mathbf{0}_{1\times rq} & N_4 & \mathbf{0}_{1\times\infty} \end{bmatrix} + N_5 \right) \boldsymbol{z}_{t-1}
\end{aligned} \tag{39}$$

where r is the number of elements in each agents' vector of idiosyncratic shocks, $\boldsymbol{v}_t(i)$. Putting it all together, we substitute equations (38) and (39) into equation (36) to arrive at

$$F = \begin{bmatrix} \begin{bmatrix} A & \mathbf{0}_{m\times\infty} \end{bmatrix} \\ K (M_1 + M_2 + M_3) + (F - KM_1) T_s - KM_2 T_{w_1} \\ K (M_1 + M_2 + M_3) + (F - KM_1) T_{w_1} - KM_2 T_{w_2} \\ K (M_1 + M_2 + M_3) + (F - KM_1) T_{w_2} - KM_2 T_{w_3} \\ \vdots \end{bmatrix} \tag{40a}$$

$$G_1 = \begin{bmatrix} P \\ KN_1 \\ KN_1 \\ KN_1 \\ \vdots \end{bmatrix} \qquad G_2 = \begin{bmatrix} \mathbf{0}_{m\times\infty} \\ \mathbf{0}_{\infty\times\infty} \\ K \begin{bmatrix} N_2 & \mathbf{0}_{1\times r} & \mathbf{0}_{1\times r} & \mathbf{0}_{1\times\infty} \end{bmatrix} \\ K \begin{bmatrix} \mathbf{0}_{1\times r} & N_2 & \mathbf{0}_{1\times r} & \mathbf{0}_{1\times\infty} \end{bmatrix} \\ \vdots \end{bmatrix} \tag{40b}$$

$$G_3 = \begin{bmatrix} \mathbf{0}_{m\times n} \\ KN_3 \\ KN_3 \\ KN_3 \\ \vdots \end{bmatrix} \qquad G_4 = \begin{bmatrix} \mathbf{0}_{m\times\infty} \\ K \left(\begin{bmatrix} N_4 & \mathbf{0}_{1\times p} & \mathbf{0}_{1\times r} & \mathbf{0}_{1\times\infty} \end{bmatrix} + N_5 \right) \\ K \left(\begin{bmatrix} \mathbf{0}_{1\times p} & N_4 & \mathbf{0}_{1\times r} & \mathbf{0}_{1\times\infty} \end{bmatrix} + N_5 \right) \\ K \left(\begin{bmatrix} \mathbf{0}_{1\times r} & \mathbf{0}_{1\times r} & N_4 & \mathbf{0}_{1\times\infty} \end{bmatrix} + N_5 \right) \\ \vdots \end{bmatrix} \tag{40c}$$

where m is the number of elements in the underlying state $(\boldsymbol{x}_t)$ and n is the number of elements in the vector of public signal noise $(\boldsymbol{e}_t)$. This confirms the conjectured structure to the law of motion and implicitly defines the coefficient matricies. Note that since the matricies in (40) are recursive, finding the solution involves finding the fixed point of the system for a given Kalman gain (K) and pre-chosen upper limit (k^*) on the number of orders of expectations to include.

5 Proof of proposition 3.

For standard problems with imperfect common knowledge, where only the hierarchy of simple-average expectations is needed,[2] an arbitrarily accurate approximation of the full solution can be achieved by selecting a cut-off, k^*, and including all orders of expectation from zero to that cut-off, provided that

1. the importance attached to higher-*order* average expectations is decreasing in the order; and
2. the unconditional variance of higher-order average expectations are bounded from above.

The first of these is imposed by assumption. In the context of the model presented here, this amounts to a restriction on the coefficients in $\boldsymbol{\lambda}_1$.[3] The second is assured by the fact that agents are rational (Bayesian) and this is common knowledge. A proof of this is provided by Nimark (2011a), although it requires one minor extension here. Since I can write $X_t = E_t(j)[X_t] + X_{t|t}^{\text{err}}(j)$ and the variance of the two sides must be equal, I have

$$Var(X_t) = Var(E_t(j)[X_t]) + Var\left(X_{t|t}^{\text{err}}(j)\right)$$

where the covariance term on the right hand side can be ignored because j's rationality implies that her expectation must be orthogonal to her expectation error. This demonstrates that

$$Var(E_t(j)[X_t]) \leq Var(X_t)$$

The Kalman filter ensures that j's expectation must have a Moving Average representation incorporating linear combinations of the complete history of all shocks that enter her signals. For a simple average of this (lemma 2 in the nimark paper), any idiosyncratic shocks will necessarily sum to zero, ensuring that the simple-average expectation must have lower variance than that of any individual agent. For weighted averages of this, the idiosyncratic shocks will not sum to zero, but the variance of the weighted-average of those shocks will be less the variance of an individual shock as shown above in corollary 1 to proposition 1. It therefore must be that

$$Var\left(\overline{E}_t[X_t]\right) \leq Var\left(\overset{1:\sim}{E}_t[X_t]\right) \leq Var\left(\overset{2:\sim}{E}_t[X_t]\right) \leq \cdots \leq Var(E_t(j)[X_t]) \leq Var(X_t)$$

The recursive structure of X_t then establishes the result.

In addition, it is *also* necessary here to define a cut-off in the number of compound expectations to include (p^*). Analogously to the cut-off in higher orders of average expectation, the researcher's ability to deliver an arbitrarily accurate approximation requires that

1. the importance attached to higher-*weighted* expectations is decreasing in the weighting; and
2. the unconditional variance of higher-weighted average expectations are bounded from above.

The first of these is implied by the fact that each (next) higher weighted average expectation enters with a (further) lag and the underlying autoregressive process ensures that agents assign decreasing importance to older signals when considering their current expectation. The second was described above and is implied directly by corollary 1 to proposition 1.

[2] That is, where there is only one compound expectation of interest ($p = 1$).
[3] See section 4 of the main article for a typical example.

6 Implementation

Implementing a finite approximation with a cut-off, p^*, in the number of weighted-averages to include still requires that the programmer take a view on how to implement the the final weight. Recall from the main text that for the simplified model with no (lagged) public signal, the law of motion is

$$\begin{aligned}
\boldsymbol{x}_t &= \rho\,\boldsymbol{x}_{t-1} && + \quad \boldsymbol{u}_t \\
\overline{E}_t\left[X_t\right] &= B\,\boldsymbol{x}_{t-1} + C\ \overline{E}_{t-1}\left[X_{t-1}\right] + D\overset{1:\sim}{E}_{t-1}\left[X_{t-1}\right] + H\boldsymbol{u}_t \\
\overset{1:\sim}{E}_t\left[X_t\right] &= B\,\boldsymbol{x}_{t-1} + C\overset{1:\sim}{E}_{t-1}\left[X_{t-1}\right] + D\overset{2:\sim}{E}_{t-1}\left[X_{t-1}\right] + H\boldsymbol{u}_t + Q\overset{1:\sim}{\boldsymbol{v}}_t \\
\overset{2:\sim}{E}_t\left[X_t\right] &= B\,\boldsymbol{x}_{t-1} + C\overset{2:\sim}{E}_{t-1}\left[X_{t-1}\right] + D\overset{3:\sim}{E}_{t-1}\left[X_{t-1}\right] + H\boldsymbol{u}_t + Q\overset{2:\sim}{\boldsymbol{v}}_t \\
&\ \ \vdots
\end{aligned}$$

$$\text{where} \quad \begin{aligned}
B &= \boldsymbol{k}_p\,\rho & H &= \boldsymbol{k}_p \\
C &= F - BS_x - DT_{w_1} & Q &= q\,\boldsymbol{k}_p \\
D &= q\,\boldsymbol{k}_s\,\boldsymbol{\lambda}_1' & &
\end{aligned}$$

with $\boldsymbol{k}_p$ being the Kalman gain applied to the private signal and $\boldsymbol{k}_s$ the Kalman gain applied to each social signal, so that the transition matrix for the full state therefore takes the following form:

$$F = \begin{bmatrix}
\rho & 0 & 0 & 0 & \cdots \\
B & C & D & 0 & \\
B & 0 & C & D & \\
B & 0 & 0 & C & \ddots \\
\vdots & & & & \ddots
\end{bmatrix}$$

For the p^{th}-weighted expectation, we have

$$\begin{aligned}
\overset{p:\sim}{E}_t\left[X_t\right] &= BS_xX_{t-1} + CT_{w_p}X_{t-1} + DT_{w_{p+1}}X_{t-1} + shocks \\
&= \boldsymbol{k}_p\,\rho S_xX_{t-1} + \left(F - \boldsymbol{k}_p\,\rho S_x - q\,\boldsymbol{k}_s\,\boldsymbol{\lambda}_1'T_{w_1}\right)T_{w_p}X_{t-1} + q\,\boldsymbol{k}_s\,\boldsymbol{\lambda}_1'T_{w_{p+1}}X_{t-1} + shocks \\
&= \left(\boldsymbol{k}_p\,\rho S_x + \left(F - \boldsymbol{k}_p\,\rho S_x\right)T_{w_p}\right)X_{t-1} + q\,\boldsymbol{k}_s\,\boldsymbol{\lambda}_1'\left(T_{w_{p+1}} - T_{w_1}T_{w_p}\right)X_{t-1} + shocks
\end{aligned}$$

When considering the expectations of agents p levels deep in the network, the component derived from consideration of agents $p+1$ levels deep is captured in the term $q\,\boldsymbol{k}_s\,\boldsymbol{\lambda}_1'\left(T_{w_{p+1}} - T_{w_1}T_{w_p}\right)X_{t-1}$. For the final weighting in the simulation, two clear possibilities are apparent:

- For the final weight, use $q\,\boldsymbol{k}_s\,\boldsymbol{\lambda}_1'\left(T_{w_{p^*}} - T_{w_1}T_{w_{p^*}}\right)X_{t-1}$
- For all weights $\Psi q\,\boldsymbol{k}_s\,\boldsymbol{\lambda}_1'\left(T_{w_{p+1}} - T_{w_1}T_{w_p}\right)X_{t-1}$ and have $\Psi = 1$ for $p < p^*$ and $\Psi = 0$ for $p = p^*$

The first option implies that agents treat competitors p and $p+1$ levels deep in the network the same, and know that all other agents take the same approach. The second option implies that agents suppose that competitors p levels deep in the network do not observe anybody so their information comes only from their public/private signals. Both options must be equivalent as $p^* \to \infty$ and, in practice, are seen to produce highly similar results.

The attached Matlab code provides an implementation of the model that uses a third alternative:

- For the final weight, use $\Psi q\, \boldsymbol{k}_s\, \boldsymbol{\lambda}_1' \left(T_{w_{p^*}} - T_{w_1} T_{w_{p^*}} \right) X_{t-1}$ where $\Psi = 1 + \epsilon$.

which assumes that agents treat competitors p and $p+1$ levels deep in the network the same *and* artificially forces them to place slightly more weight on them when constructing the Kalman filter in order to crudely capture the unsimulated higher-weighted expectations. Doing this improves the implementation's robustness to numerical instability and allows simulations with higher numbers of observees (q).

Numerical instability

Although equations (35) and (40) provide the algorithm through which to iterate, as written they are extremely memory intensive and prone to numerical instability. This problem worsens as q increases and, for moderate-to-high persistence in the underlying state, the solution can only be found for very low values of q.

Without recourse to standard UD-factorisation techniques (see the main text), then in addition to the avoidance of stacking the state vector already implemented and the implementation of p^* mentioned above, I also deploy the following techniques to improve the algorithm's performance:

Avoid unnecessary iteration

As mentioned above, the network learning problem involves finding convergent solutions to the filter and the law of motion, each taking the other as given. In principle, the fixed point may therefore be found by finding the convergent result of one within each iteration of the other – for example:

```
repeat
    Update the filter by one step using equation (35)
    repeat
        Update the law of motion by one step using equation (40)
    until the law of motion converges
until the filter converges
```

This set-up is $O(n^2)$, however, even before examining the complexity of the one-step processes, and in practice is more likely to suffer from numerical stability issues. Instead, for a given set of signals, I find the fixed point by updating the filter and the law of motion incrementally within the same loop:

```
repeat
    Update the filter by one step using equation (35)
    Update the law of motion by one step using equation (40)
until both the filter and the law of motion converge
```

Avoid temporary creation of unnecessarily large matrices

The solution as presented above (see equations 35a and 35b) involves the temporary creation (and multiplication) of matrices that are $(2+q) \times N$ square, where N is the size of X_t and q is the number of other agents observed.

The implementation presented in the attached Matlab code keeps the public/private signals and the social signals separate (i.e. it breaks the M_* and N_* matrices into their constituent components) to avoid this and to exploit the fact that each social signal will be treated identically.

Pay close attention to operation order

Because matrix addition and subtraction are of order $O\left(n^2\right)$ while (naive) matrix multiplication and inversion are of order $O\left(n^3\right)$, the order in which expressions are calculated can affect the number of operations required.

For example, although mathematically equivalent, the computational complexity of calculating $(A+B)\times C$ is *less* than that of $(A\times C)+(B\times C)$ because the former involves only a single multiplication.

7 Extending the model to dynamic actions

We here consider an illustrative example of extending the model of this chapter to consideration of dynamic actions. In particular, we allow agents' decision rules to be slightly more general, with an inclusion of agents' expectations regarding the next-period average action. That is, we suppose that individual decisions are made according to the following rule:

$$g_t(i) = \boldsymbol{\alpha}' \boldsymbol{s}_t^p(i) + \boldsymbol{\eta}_x' E_t(i)\left[X_t\right] + \eta_y E_t(i)\left[\overline{g}_t\right] + \eta_z E_t(i)\left[\overline{g}_{t+1}\right] \tag{41}$$

where agents' private signals are formed as

$$\boldsymbol{s}_t^p(i) = B\boldsymbol{x}_t + Q\boldsymbol{v}_t(i)$$

We retain the assumption that the underlying state follows an AR(1) process:

$$\boldsymbol{x}_t = A\boldsymbol{x}_{t-1} + P\boldsymbol{u}_t$$

and still suppose that the full hierarchy of expectations regarding the underlying state is given by:

$$X_t = \mathbb{E}_t^{(0:\infty)}\left[\boldsymbol{x}_t\right]$$

Our goal is to show that $g_t(i)$ may be expressed in the general form

$$g_t(i) = \boldsymbol{\lambda}_0' w_{t-1} + \boldsymbol{\lambda}_2' X_t + \boldsymbol{\lambda}_1' E_t(i)\left[X_t\right] + \boldsymbol{\lambda}_3' \boldsymbol{v}_t(i)$$

To do this, we start by taking the simple average of equation (41) to give:

$$\overline{g}_t = \boldsymbol{\alpha}' B\boldsymbol{x}_t + \boldsymbol{\eta}_x' \overline{E}_t\left[X_t\right] + \eta_y \overline{E}_t\left[\overline{g}_t\right] + \eta_z \overline{E}_t\left[\overline{g}_{t+1}\right]$$

To keep the notation clean, define $\theta_t \equiv \boldsymbol{\alpha}' B\boldsymbol{x}_t + \boldsymbol{\eta}_x' \overline{E}_t\left[X_t\right]$ so that

$$\overline{g}_t = \theta_t + \eta_y \overline{E}_t\left[\overline{g}_t\right] + \eta_z \overline{E}_t\left[\overline{g}_{t+1}\right]$$

We now substitute this equation back into itself in the second element ($\eta_y \overline{E}_t\left[\overline{g}_t\right]$):

$$\overline{g}_t = \theta_t + \eta_y \overline{E}_t\left[\theta_t\right] + \eta_y^2 \overline{E}_t^{(2)}\left[\overline{g}_t\right] + \eta_z \overline{E}_t\left[\overline{g}_{t+1}\right] + \eta_y \eta_z \overline{E}_t^{(2)}\left[\overline{g}_{t+1}\right]$$

Repeating this process, in the limit (and using the fact that $\eta_y \in (0,1)$ and assuming that average expectations do not explode), this becomes:

$$\overline{g}_t = \left(\sum_{k=0}^{\infty} \eta_y^k \overline{E}_t^{(k)}\left[\theta_t\right]\right) + \left(\eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)}\left[\overline{g}_{t+1}\right]\right)$$

now briefly consider θ_t and simple-average expectations of θ_t. We can write that:

$$\begin{aligned}
\theta_t &= \boldsymbol{\alpha}' B \boldsymbol{x}_t + \boldsymbol{\eta}_x' \overline{E}_t^{(1)}\left[X_t\right] \\
\overline{E}_t^{(1)}\left[\theta_t\right] &= \boldsymbol{\alpha}' B \overline{E}_t^{(1)}\left[\boldsymbol{x}_t\right] + \boldsymbol{\eta}_x' \overline{E}_t^{(2)}\left[X_t\right] \\
\overline{E}_t^{(2)}\left[\theta_t\right] &= \boldsymbol{\alpha}' B \overline{E}_t^{(2)}\left[\boldsymbol{x}_t\right] + \boldsymbol{\eta}_x' \overline{E}_t^{(3)}\left[X_t\right] \\
&\ldots
\end{aligned}$$

next, suppose that the matrix T_s selects the simple-average expectation of X_t from X_t:

$$\overline{E}_t^{(1)}\left[X_t\right] = T_s X_t$$

and that the matrix S selects $\boldsymbol{x}_t$ from X_t (obviously $S = \begin{bmatrix} I_l & 0_{l\times\infty}\end{bmatrix}$ where l is the number of elements in $\boldsymbol{x}_t$):

$$\boldsymbol{x}_t = S X_t$$

Then we can write:

$$\begin{aligned}
\theta_t &= \left(\boldsymbol{\alpha}' BS + \boldsymbol{\eta}_x' T_s\right) X_t \\
\overline{E}_t^{(1)}\left[\theta_t\right] &= \left(\boldsymbol{\alpha}' BS + \boldsymbol{\eta}_x' T_s\right) T_s X_t \\
\overline{E}_t^{(2)}\left[\theta_t\right] &= \left(\boldsymbol{\alpha}' BS + \boldsymbol{\eta}_x' T_s\right) T_s^2 X_t \\
&\ldots
\end{aligned}$$

or, in general,

$$\overline{E}_t^{(k)}\left[\theta_t\right] = \left(\boldsymbol{\alpha}' BS + \boldsymbol{\eta}_x' T_s\right) T_s^k X_t$$

The average period-t action can therefore be written as

$$\begin{aligned}
\overline{g}_t &= \left(\boldsymbol{\alpha}' BS + \boldsymbol{\eta}_x' T_s\right)\left(\sum_{k=0}^{\infty}\left(\eta_y T_s\right)^k\right) X_t + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)}\left[\overline{g}_{t+1}\right] \\
&= \left(\boldsymbol{\alpha}' BS + \boldsymbol{\eta}_x' T_s\right)\left(I - \eta_y T_s\right)^{-1} X_t + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)}\left[\overline{g}_{t+1}\right] \\
&= \boldsymbol{\beta}' X_t + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)}\left[\overline{g}_{t+1}\right]
\end{aligned}$$

where $\boldsymbol{\beta}' \equiv \left(\boldsymbol{\alpha}' BS + \boldsymbol{\eta}_x' T_s\right)\left(I - \eta_y T_s\right)^{-1}$. Next, substitute this back into itself for the next-period average action:

$$\begin{aligned}
\overline{g}_t &= \boldsymbol{\beta}' X_t + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)}\left[\boldsymbol{\beta}' X_{t+1} + \eta_z \sum_{l=1}^{\infty} \eta_y^{l-1} \overline{E}_{t+1}^{(l)}\left[\overline{g}_{t+2}\right]\right] \\
&= \boldsymbol{\beta}' X_t + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \boldsymbol{\beta}' \overline{E}_t^{(k)}\left[X_{t+1}\right] + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)}\left[\eta_z \sum_{l=1}^{\infty} \eta_y^{l-1} \overline{E}_{t+1}^{(l)}\left[\overline{g}_{t+2}\right]\right]
\end{aligned}$$

next, we use the following conjectured aspect of the law of motion for X_t:

$$E_t(i)\left[X_{t+1}\right] = E_t(i)\left[F X_t\right]$$

for some matrix of parameters F. This implies that

$$\overline{E}_t^{(k)}\left[X_{t+1}\right] = F \overline{E}_t^{(k)}\left[X_t\right]$$

and hence that

$$\begin{aligned}
\overline{g}_t &= \boldsymbol{\beta}' X_t + \eta_z \boldsymbol{\beta}' F \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)} \left[X_t\right] + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)} \left[\eta_z \sum_{l=1}^{\infty} \eta_y^{l-1} \overline{E}_{t+1}^{(l)} \left[\overline{g}_{t+2}\right]\right] \\
&= \boldsymbol{\beta}' X_t + \eta_z \boldsymbol{\beta}' F \left(\sum_{k=1}^{\infty} \eta_y^{k-1} T_s^k\right) X_t + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)} \left[\eta_z \sum_{l=1}^{\infty} \eta_y^{l-1} \overline{E}_{t+1}^{(l)} \left[\overline{g}_{t+2}\right]\right] \\
&= \boldsymbol{\beta}' X_t + \eta_z \boldsymbol{\beta}' F T_s \left(I - \eta_y T_s\right)^{-1} X_t + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)} \left[\eta_z \sum_{l=1}^{\infty} \eta_y^{l-1} \overline{E}_{t+1}^{(l)} \left[\overline{g}_{t+2}\right]\right]
\end{aligned}$$

next, expand the $\overline{g}_{t+2}$ term to give

$$\begin{aligned}
\overline{g}_t &= \boldsymbol{\beta}' X_t + \eta_z \boldsymbol{\beta}' F T_s \left(I - \eta_y T_s\right)^{-1} X_t \\
&\quad + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)} \left[\eta_z \sum_{l=1}^{\infty} \eta_y^{l-1} \overline{E}_{t+1}^{(l)} \left[\boldsymbol{\beta}' X_{t+2} + \eta_z \sum_{m=1}^{\infty} \eta_y^{m-1} \overline{E}_{t+2}^{(m)} \left[\overline{g}_{t+3}\right]\right]\right] \\
&= \boldsymbol{\beta}' X_t \\
&\quad + \eta_z \boldsymbol{\beta}' F T_s \left(I - \eta_y T_s\right)^{-1} X_t \\
&\quad + \boldsymbol{\beta}' \left(\eta_z F T_s \left(I - \eta_y T_s\right)^{-1}\right)^2 X_t \\
&\quad + \eta_z \sum_{k=1}^{\infty} \eta_y^{k-1} \overline{E}_t^{(k)} \left[\eta_z \sum_{l=1}^{\infty} \eta_y^{l-1} \overline{E}_{t+1}^{(l)} \left[\eta_z \sum_{m=1}^{\infty} \eta_y^{m-1} \overline{E}_{t+2}^{(m)} \left[\overline{g}_{t+3}\right]\right]\right]
\end{aligned}$$

Continued substitution then arrives at:

$$\overline{g}_t = \boldsymbol{\beta}' \sum_{j=0}^{\infty} \left(\eta_z F T_s \left(I - \eta_y T_s\right)^{-1}\right)^j X_t$$

which, in turn, becomes

$$\overline{g}_t = \underbrace{\left(\boldsymbol{\alpha}' B S + \boldsymbol{\eta}_x' T_s\right) \left(I - \eta_y T_s\right)^{-1} \left(I - \eta_z F T_s \left(I - \eta_y T_s\right)^{-1}\right)^{-1}}_{\equiv \boldsymbol{a}'} X_t$$

Using this simple expression of $\overline{g}_t = \boldsymbol{a}' X_t$, we can substitute it back into the agents' individual decision rule to obtain

$$\begin{aligned}
g_t(i) &= \boldsymbol{\alpha}' \left(B \boldsymbol{x}_t + Q \boldsymbol{v}_t(i)\right) + \left(\boldsymbol{\eta}_x' + \eta_y \boldsymbol{a}' + \eta_z \boldsymbol{a}' F\right) E_t(i) \left[X_t\right] \\
&= \underbrace{\boldsymbol{\alpha}' B}_{\boldsymbol{\lambda}_2'} \boldsymbol{x}_t + \underbrace{\left(\boldsymbol{\eta}_x' + \eta_y \boldsymbol{a}' + \eta_z \boldsymbol{a}' F\right)}_{\gamma_3'} E_t(i) \left[X_t\right] + \underbrace{\boldsymbol{\alpha}' Q}_{\gamma_4'} \boldsymbol{v}_t(i)
\end{aligned}$$

which is now in the necessary form. As an aside, taking a simple average of this gives

$$\overline{g}_t = \boldsymbol{\alpha}' B S X_t + \left(\boldsymbol{\eta}_x' + \eta_y \boldsymbol{a}' + \eta_z \boldsymbol{a}' F\right) \overline{E}_t \left[X_t\right]$$

which implies the following constraint on the coefficients of the decision rule ($\boldsymbol{\alpha}$, $\boldsymbol{\eta}_x$, $\boldsymbol{\eta}_y$, $\boldsymbol{\eta}_z$) and the expectation transition matrix (F):

$$\boldsymbol{a}' = \boldsymbol{\alpha}' B S + \left(\boldsymbol{\eta}_x' + \eta_y \boldsymbol{a}' + \eta_z \boldsymbol{a}' F\right) T_s$$

References

Nimark, K. (2011a): "Dynamic Higher Order Expectations," Universitat Pompeu Fabra Economics Working Papers No 1118.

——— (2011b): "A low dimensional Kalman Filter for systems with lagged observables," mimeo.