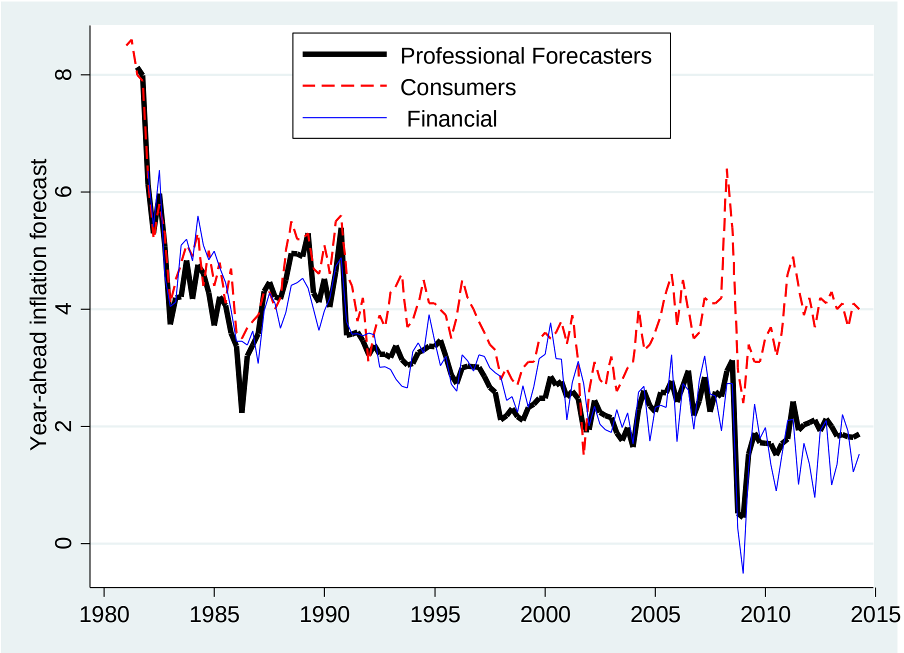
ONLINE APPENDIX

Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts

Olivier Coibion
U.T. Austin and NBER

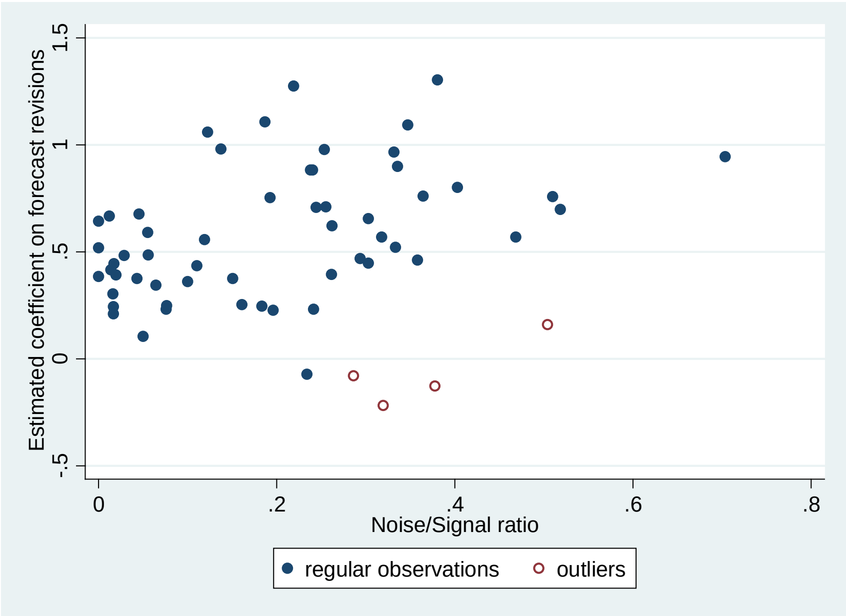
Yuriy Gorodnichenko
U.C. Berkeley and NBER

Appendix Figure 1: Inflation Forecasts from Professional Forecasters, Consumers and Financial Markets



Note: The figure plots the one-year ahead CPI forecasts from the Survey of Professional Forecasters, the Michigan Survey of Consumers, and financial markets. See section 3.1 for details.

Appendix Figure 2: Noise-Signal Ratios and Estimated Coefficients on Forecast Revisions



Note: The table plots the noise/signal ratio for each country/variable pair (horizontal axis) where noise is measured using the size of revisions to the data, as discussed in section 3.3. The vertical axis indicates the coefficient on forecast revisions from estimating (19) for each country/macroeconomic variable pair. The empty circles are outliers as identified by robust S-regression of (21) in the text.

Appendix Table 1. Properties of inflation forecasts

	Panel A: Comparison of Mean Squared Error (MSE)		
	Survey of Professional Forecasters (SPF)	Michigan Survey of Consumers (MSC)	Financial markets (FIN)
	(1)	(2)	(3)
MSE	1.146	2.690	1.383
	(0.191)	(0.505)	(0.191)
p-value of equality	-	<0.01	0.003
Observations	112	112	112

	Panel B: Predictability of ex-post CPI inflation $\pi_{t+1,t+4}$		
	(1)	(2)	(3)
SPF, $F_t\pi_{t+1,t+4}$	0.814***	0.717**	0.955***
	(0.247)	(0.303)	(0.356)
MSC, $F_t\pi_{t+1,t+4}$	-0.240		-0.267
	(0.326)		(0.319)
FIN, $F_t\pi_{t+1,t+4}$		-0.040	-0.123
		(0.250)	(0.217)
R-squared	0.358	0.347	0.359
Observations	127	127	127

Notes: In Panel A, figures in parentheses are the standard errors of the MSE estimates. The last row in Panel A reports the *p*-value of the *t*-test of equality of MSE for SPF and an alternative source of forecasts. In Panel B, standard errors are in parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

Appendix Table 2. Pooled Estimates of the Expectations Formation Process

Dependent variable Forecast error $x_{t+h} - F_t x_{t+h}$	U.S. SPF 1968-2014 5 Variables			U.S. SPF 1982-2014 13 Variables			Cross-Country Professional Forecasters 1989-2010 5 Variables 12 countries		
	OLS	FE	FE + time dummies	OLS	FE	FE + time dummies	OLS	FE	FE + time dummies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Forecast revision	0.395**	0.391**	0.313*	0.661***	0.661***	0.639***	0.697***	0.642***	0.519***
$(F_t x_{t+h} - F_{t-1} x_{t+h})$	(0.184)	(0.181)	(0.115)	(0.184)	(0.184)	(0.134)	(0.143)	(0.139)	(0.072)
Observations	3,577	3,577	3,577	6,654	6,654	6,654	22,340	22,340	22,340
R-squared	0.019	0.018	0.131	0.030	0.030	0.079	0.048	0.043	0.234
Panel B									
$F_t x_{t+h}$	0.428**	0.435**	0.362**	0.662***	0.684***	0.645***	0.713***	0.660***	0.522***
	(0.179)	(0.177)	(0.108)	(0.179)	(0.178)	(0.156)	(0.140)	(0.131)	(0.112)
$F_{t-1} x_{t+h}$	-0.478**	-0.520***	-0.467***	-0.617***	-0.621***	-0.620**	-0.775***	-0.734***	-0.538***
	(0.174)	(0.181)	(0.072)	(0.185)	(0.201)	(0.209)	(0.119)	(0.097)	(0.085)
p-value $(\beta + \gamma) = 0$	0.167	0.298	0.265	0.499	0.564	0.794	0.065	0.229	0.757
Observations	3,575	3,575	3,575	6,654	6,649	6,649	22,340	22,340	22,340
R-squared	0.024	0.022	0.136	0.031	0.033	0.081	0.050	0.046	0.234

Notes: The table reports estimates of specification (10) in Panels A and B respectively. Driscoll-Kraay (1998) standard errors are in parentheses in columns (1), (2), (4), (5), (7), (8). Robust standard errors clustered by forecasted variable are in parentheses in columns (3), (6) and (9). Fixed effects in columns (2), (3), (5), (6), (8) and (9) are for each combination of country, variable, and forecast horizon. Time dummies in columns (3), (6) and (9) are for each time period (calendar quarter). ***, **, * denote significance at 0.01, 0.05, and 0.10 levels.

Appendix A: Bias in OLS Estimates under Common Noise

Fundamental/State equation: $\pi_t = \rho\pi_{t-1} + v_t$ where $v_t \sim iid\ N(0, \Sigma_v)$
Measurement equation: $y_t = \pi_t + e_t + \omega_{it}$ where $\omega_{it} \sim iid\ N(0, \Sigma_\omega)$ is the idiosyncratic noise, $e_t \sim iid\ N(0, \Sigma_e)$ is the common noise shock. We assume that all shocks are uncorrelated contemporaneously and at all leads and lads.

Denote the one step-ahead forecast error for the forecast $\pi_{t|t-1}(i) \equiv E(\pi_t | y_{i,t-1}, y_{i,t-2}, \dots)$ in the Kalman filter with

$$\Psi \equiv \Sigma_\pi(t|t-1) \equiv E\left(\left(\pi_{t|t-1}(i) - \pi_t\right)\left(\pi_{t|t-1}(i) - \pi_t\right)'\right).$$

We can find Ψ from the Ricatti equation $\Psi = \rho^2\{\Psi - \Psi(\Psi + \Sigma_\omega + \Sigma_e)^{-1}\Psi\} + \Sigma_v$. Denote the gain of the Kalman filter with $G = \Psi(\Psi + \Sigma_\omega + \Sigma_e)^{-1}$. The forecast for the unobserved state π_t evolves as follows:

$$\pi_{t|t}(i) = \pi_{t|t-1}(i) + G\left(y_t - y_{t|t-1}(i)\right) \tag{A.1}$$

with $y_{t|t-1}(i) = \pi_{t|t-1}(i)$. Let $\overline{\pi_{t+h|t}} \equiv E_i\left(\pi_{t+h|t}(i)\right)$ be the average forecast across agents. Using $E_i(\omega_{it}) = 0$, we can re-write equation (A.1) as

$$\left(\pi_{t+h} - \overline{\pi_{t+h|t}}\right) = \frac{1-G}{G}\left(\overline{\pi_{t+h|t}} - \overline{\pi_{t+h|t-1}}\right) - e_t + v_{t+h,t},$$

where $v_{t+h,t}$ is a linear combination of shocks v_t occurring between t and $t+h$ which constitutes a rational expectations error.

Suppose one runs an OLS regression of the following form

$$\left(\pi_{t+h} - \overline{\pi_{t+h|t}}\right) = \beta\left(\overline{\pi_{t+h|t}} - \overline{\pi_{t+h|t-1}}\right) + error_t.$$

Then, using $E\left(\overline{\pi_{t|t-1}}e_t\right) = 0$ and $E(\pi_t e_t) = 0$, we obtain

$$\begin{aligned}
\hat{\beta}^{OLS} &= \frac{E\left(\left\{\pi_{t+h} - \overline{\pi_{t+h|t}}\right\}\left\{\overline{\pi_{t+h|t}} - \overline{\pi_{t+h|t-1}}\right\}\right)}{E\left(\left\{\overline{\pi_{t+h|t}} - \overline{\pi_{t+h|t-1}}\right\}^2\right)} = \frac{1-G}{G} + \frac{-E\left(\left\{\overline{\pi_{t+h|t}} - \overline{\pi_{t+h|t-1}}\right\}e_t\right)}{E\left(\left\{\overline{\pi_{t+h|t}} - \overline{\pi_{t+h|t-1}}\right\}^2\right)} \\
&= \frac{1-G}{G} + \frac{-E\left(\overline{\pi_{t+h|t}}e_t\right)}{E\left(\left\{\overline{\pi_{t+h|t}} - \overline{\pi_{t+h|t-1}}\right\}^2\right)} = \frac{1-G}{G} + \frac{-E\left(\rho^h\left\{(1-G)\overline{\pi_{t|t-1}} + G\pi_t + Ge_t\right\}e_t\right)}{E\left(\left\{\overline{\pi_{t+h|t}} - \overline{\pi_{t+h|t-1}}\right\}^2\right)} \\
&= \frac{1-G}{G} - \frac{\rho^h G\Sigma_e}{E\left(\left\{\overline{\pi_{t+h|t}} - \overline{\pi_{t+h|t-1}}\right\}^2\right)}.
\end{aligned}$$

Hence, $\hat{\beta}^{OLS}$ is biased downward and if we find $\hat{\beta}^{OLS} > 0$ then the true $\beta = \frac{1-G}{G}$ must be even larger.

Appendix B: Noisy-Information Model under Generalized Dynamics

Consider the following state-space representation of the AR(p) inflation process (i.e., the fundamental) and observed signal about inflation:

State: $z_t \equiv \begin{bmatrix} \pi_t \\ \vdots \\ \pi_{t-p+1} \end{bmatrix} = \begin{bmatrix} b_1 & \dots & \dots & b_p \\ 1 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 1 & 0 \end{bmatrix} z_{t-1} + S'v_t = Bz_{t-1} + S'v_t$ (B.1)

where $S = [1\ 0\ \dots\ 0]$, $v_t \sim iid\ N(0, \sigma_v^2)$ is a shock to fundamental with $\Sigma_v = \sigma_v^2 S'S$.

Measurement: $y_{it} = Hz_t + \omega_{it}$ where $H = [1\ 0\ \dots\ 0]$, $\omega_{it} \sim iid\ N(0, \Sigma_\omega)$ is the agent specific shock which is uncorrelated across agents. We assume that $E(v_t \omega_{it}') = 0$, that is shocks to fundamentals v_t and measurement error shocks ω_{it} are independent.

Denote the one step-ahead forecast error for the forecast $z_{t|t-1}(i) \equiv E(z_t | y_{i,t-1}, y_{i,t-2}, \dots)$ in the Kalman filter with $\Psi \equiv \Sigma_z(\mathrm{t|t-1}) \equiv E\left(\left(z_{t|t-1}(i) - z_t\right)\left(z_{t|t-1}(i) - z_t\right)'\right)$. We can find Ψ from the Ricatti equation: $\Psi = B\{\Psi - \Psi H'(H\Psi H' + \Sigma_\omega)^{-1} H\Psi\}B' + \Sigma_v$. Denote the gain of the Kalman filter with $G = \Psi H'(H\Psi H' + \Sigma_\omega)^{-1}$.

The forecast for the unobserved state z_t evolves as follows:

$$z_{t|t}(i) = z_{t|t-1}(i) + G\left(y_{it} - y_{t|t-1}(i)\right) = z_{t|t-1}(i) + G\left(Hz_t + \omega_{it} - Hz_{t|t-1}(i)\right) \quad \text{(B.2)}$$

where $y_{t|t-1}(i) = Hz_{t|t-1}(i)$. After taking averages across agents (and hence dropping index i) and using $Hz_t = \pi_t$, $Hz_{t|t-1} = \pi_{t|t-1}$, and $Hz_{t|t} = \pi_{t|t}$, we note that

$$\pi_{t+h} - \pi_{t+h|t} = H\left(z_{t+h} - z_{t+h|t}\right) = HB^h\left(z_t - z_{t|t}\right) + REerror \quad \text{(B.3)}$$

After re-arranging (B.2) averaged across agents and using generalized inverses (denoted as "+" in the power), we obtain

$$\pi_{t+h} - \pi_{t+h|t} = H\beta\left(z_{t+h|t} - z_{t+h|t-1}\right) + REerror \quad \text{(B.4)}$$

where $\beta \equiv B^h\{GH\}^+(I - GH)\left(B^h\right)^{-1} = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots \\ \beta_{21} & \dots & \\ \dots & & \beta_{pp} \end{bmatrix}$ and $H\beta = [\beta_{11}\ \ \dots\ \ \beta_{1p}]$. As a result,

$$\pi_{t+h} - \pi_{t+h|t} = \beta_{11}\left(\pi_{t+h|t} - \pi_{t+h|t-1}\right) + \beta_{12}\left(\pi_{t+h-1|t} - \pi_{t+h-1|t-1}\right) + \cdots$$
$$+\beta_{1p}\left(\pi_{t+h-(p-1)|t} - \pi_{t+h-(p-1)|t-1}\right) + REerror \quad \text{(B.5)}$$

Note that (B.5) nests the AR(1) case in which $\beta_{11} = (1 - G)/G$. The only difference in (B.4) relative to the AR(1) case is that we have $(I - GH)$ as a *matrix* (rather than a scalar) measure of information rigidity but there is no difference in interpretation. In summary, the forecast error depends on how agents revise their forecast of inflation that period as well as previous *p* periods.

Expression in (B.5) highlights the challenge of measuring information rigidity: forecast errors and revisions capture both information rigidity and time series properties of the data generating process. However, one can recover information rigidity in a special case. Note that the gain of the Kalman filter in this case is given by $G = [G_1\ \ \cdots\ \ G_p]'$. It follows that

$$GH = \begin{bmatrix} G_1 \\ \vdots \\ G_p \end{bmatrix} [1\ \ 0\ \ \cdots\ \ 0] = \begin{bmatrix} G_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ G_p & 0 & \cdots & 0 \end{bmatrix} \quad \text{(B.6)}$$

Also, the generalized inverse for GH is given by

$$\{GH\}^{+}=\begin{bmatrix}\frac{G_1}{(G_1+G_2+\cdots+G_p)^2} & \cdots & \frac{G_p}{(G_1+G_2+\cdots+G_p)^2}\\ 0 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 0\end{bmatrix} \tag{B.7}$$

Hence,

$$\{GH\}^{+}(I-GH)=\begin{bmatrix}\frac{G_1}{(G_1+G_2+\cdots+G_p)^2}-1 & \frac{G_2}{(G_1+G_2+\cdots+G_p)^2} & \cdots & \frac{G_p}{(G_1+G_2+\cdots+G_p)^2}\\ 0 & & \cdots & 0\\ \vdots & & \ddots & \vdots\\ 0 & & \cdots & 0\end{bmatrix} \tag{B.8}$$

Therefore, if we consider specification (B.5) for $h=0$, we obtain

$$\begin{aligned}\pi_t-\pi_{t|t}=&\left\{\frac{G_1}{(G_1+G_2+\cdots+G_p)^2}-1\right\}(\pi_{t|t}-\pi_{t|t-1})\\ &+\left\{\frac{G_2}{(G_1+G_2+\cdots+G_p)^2}\right\}(\pi_{t-1|t}-\pi_{t-1|t-1})+\cdots\\ &+\left\{\frac{G_p}{(G_1+G_2+\cdots+G_p)^2}\right\}(\pi_{t-(p-1)|t}-\pi_{t-(p-1)|t-1})+REerror.\end{aligned} \tag{B.9}$$

The key advantage of using $h=0$ is that B^h and $(B^h)^{-1}$ drop out from the expression for β. As a result, one can read coefficients in the Kalman gain G straight off the estimated coefficients in equation (B.9).

Now we consider how these results generalize to VAR(p) settings. Under VAR dynamics, the relevant forecast revisions on the right-hand side of equation (B.5) also include revisions in forecasts for other variables included in the VAR. As an example consider VAR(1) process (i.e., the fundamental is VAR(1)) with inflation being the variable of interest:

State: $$z_t\equiv\begin{bmatrix}\pi_t\\ x_t^{(1)}\\ \vdots\\ x_t^{(k-1)}\end{bmatrix}=\begin{bmatrix}b_{11} & \ldots & b_{1k}\\ \vdots & \ddots & \vdots\\ b_{k1} & \ldots & b_{kk}\end{bmatrix}z_{t-1}+v_t=Bz_{t-1}+v_t \tag{B.10}$$

where $x_t^{(1)},\ldots,x_t^{(k-1)}$ are macroeconomic variables other than inflation, $v_t\sim iid\ N(0,\Sigma_v)$ is a vector shock to fundamentals.

Measurement: $y_{it}=z_t+\omega_{it}$ where $\omega_{it}\sim iid\ N(0,\Sigma_\omega)$ with $E(v_t\omega_{it}')=0$. Given that $Hz_t=\pi_t$, $Hz_{t|t-1}=\pi_{t|t-1}$, and $Hz_{t|t}=\pi_{t|t}$, we immediately reproduce (B.3) and (B.4). Consequently,

$$\begin{aligned}\pi_{t+h}-\pi_{t+h|t}=&\beta_{11}(\pi_{t+h|t}-\pi_{t+h|t-1})+\beta_{12}\left(x_{t+h|t}^{(1)}-x_{t+h|t-1}^{(1)}\right)+\cdots\\ &+\beta_{1k}\left(x_{t+h|t}^{(k-1)}-x_{t+h|t-1}^{(k-1)}\right)+REerror.\end{aligned} \tag{B.11}$$

It follows that VAR(p) equivalent of (B.5) is

$$\begin{aligned}\pi_{t+h}-\pi_{t+h|t}=&\textstyle\sum_{s=0}^{p-1}\beta_{s1}(\pi_{t+h-s|t}-\pi_{t+h-s|t-1})\\ &+\textstyle\sum_{s=0}^{p-1}\sum_{m=1}^{k-1}\beta_{s,m+1}\left(x_{t+h-s|t}^{(m)}-x_{t+h-s|t-1}^{(m)}\right)+REerror.\end{aligned} \tag{B.12}$$

The next question is how to measure the degree of information rigidity when there are multiple variables in the state equation. Consider the following approach. In the Kalman filter, we make a one-step ahead forecast for the unobserved state variables ($z_{t|t-1}$) and then revise this forecast in light of information received in the next period ($z_{t|t}$). Because we have more information at time t, the forecast error ($z_t - z_{t|t}$) is smaller than the forecast error ($z_t - z_{t|t-1}$) in the sense that

$$\Omega \equiv \text{var}\left(z_t - z_{t|t}\right) < \text{var}\left(z_t - z_{t|t-1}\right) \equiv \Psi. \tag{B.13}$$

If signals received at time t are informative and, hence, we put a lot of weight on these signals, the reduction in the forecast error is large. If signals are not informative, we put a lot of weight on our prior beliefs and hence the size of the forecast error changes only a little. In other words, the change in the size of forecast error is informative about the degree of information frictions. To see this formally, we can use recursions of the Kalman filter to relate Ω and Ψ:

$$\Omega = \Psi - G\{\Psi + \Sigma_\omega\}G'. \tag{B.14}$$

Matrix $G\{\Psi + \Sigma_\omega\}G'$ measures by how much forecast errors shrink relative to Ψ. Using the expression for the Kalman gain G, we obtain

$$\begin{aligned}\Omega &= \Psi - G\{\Psi + \Sigma_\omega\}G' \\ &= \Psi - G\{\Psi + \Sigma_\omega\}\{\Psi(\Psi + \Sigma_\omega)^{-1}\}' \quad \text{(use the definition of } G\text{)} \\ &= \Psi - G\{\Psi + \Sigma_\omega\}(\Psi + \Sigma_\omega)^{-1}\Psi' \\ &= \Psi - G\Psi' \\ &= \Psi - G\Psi \quad \text{(}\Psi\text{ is symmetric, it's a covariance)} \\ &= \Psi(I - G) = (I - G)\Psi.\end{aligned}$$

Let $\Psi = \begin{bmatrix}\psi_{11} & \cdots & \psi_{1p} \\ \vdots & \ddots & \vdots \\ \psi_{p1} & \cdots & \psi_{pp}\end{bmatrix}$ where, for example, ψ_{11} is the forecast error for inflation π_t when the forecast is made at time $t-1$. Consider what happens with each element when we move from Ψ to Ω. It follows that

$$(I - G)\Psi = \begin{bmatrix}(1 - G_{11})\psi_{11} & \cdots & \left(1 - G_{1p}\right)\psi_{1p} \\ \vdots & \ddots & \vdots \\ \left(1 - G_{p1}\right)\psi_{p1} & \cdots & \left(1 - G_{pp}\right)\psi_{pp}\end{bmatrix} \tag{B.15}$$

So the variance of the forecast error for inflation π_t changes from ψ_{11} to $(1 - G_{11})\psi_{11}$. Dividing one variance by another, we obtain that the variance of forecast errors for inflation is reduced by a factor of $\frac{\Omega_{1,1}}{\Psi_{1,1}} = (1 - G_{11})$. This factor summarizes the amount of information received at time t when signal y_t is observed. Likewise, one can use $\frac{\Omega_{j,j}}{\Psi_{j,j}} = (1 - G_{jj})$ to measure information rigidity for the j^{th} variables in the VAR.

This result has an intuitive mapping to the univariate AR(1) case discussed in the paper, in the sense that $(1 - G_{11})$ measures the weight one assigns on prior information and thus measures the degree of information rigidity. If G_{11} is close to one (the signal is extremely informative (practically reveals the state variable), we put a lot weight on new information). In the AR(p) case, the first element of G has a similar interpretation.

The next question is how one can recover diagonal elements of matrix G. In the AR(p) case, G is a vector. In the VAR(1) case, G is a $k \times k$ matrix. The VAR(1) analogue of equation (B.8) is $\{G\}^{+}(I - G) = \Sigma_\omega \Psi^{-1}$ which is a complicated expression in general and thus one cannot recover, for example, G_{11} from a univariate regression at horizon $h = 0$. However, one can recover G from a multivariate regression of the following form:

$$z_t - z_{t|t} = M\left(z_{t|t} - z_{t|t-1}\right) + \ error \tag{B.16}$$

which generalizes equation (B.12) to include all variables in VAR(1) rather than focus on just one variable in the VAR. We know from equation (B.4) that

$$M = G^{+}(I - G) \quad \text{(B.17)}$$

Since we can estimate M, we can recover G:

$$M = G^{+}(I - G) \Rightarrow GM = I - G \Rightarrow G(M + I) = I \Rightarrow G = (M + I)^{-1}. \quad \text{(B.18)}$$

Appendix C: Noisy-Information Model with Heterogeneous Signal-Noise Ratios

For simplicity, assume a random walk process for x s.t. $\rho = 1$. Assume Kalman gains are distributed normally across population with cross-sectional variance σ_G^2 and mean G. Then Coibion and Gorodnichenko (2012) show that

$$x_{t|t} = -\sigma_G^2 \left\{ \sum_{k=0}^{\infty} A_k(G) x_{t-k-1} \right\} + (1-G) x_{t-1|t-1} + G x_t$$

where $x_{t+h,t}$ is the average forecast of x_{t+h} made at time *t* across the population and where *G* is the average gain. This can be rewritten as

$$x_{t+1|t} = -\sigma_G^2 \left\{ \sum_{k=0}^{\infty} A_k(G) x_{t-k-1} \right\} + (1-G) x_{t+1|t-1} + G x_t$$

$$\Rightarrow x_{t+1|t} = -\sigma_G^2 \left\{ \sum_{k=0}^{\infty} A_k(G) x_{t-k-1} \right\} + (1-G) x_{t+1|t-1} + G(x_{t+1} - v_{t+1})$$

$$\Rightarrow G\left(x_{t+1} - x_{t+1|t}\right) = \sigma_G^2 \left\{ \sum_{k=0}^{\infty} A_k(G) x_{t-k-1} \right\} + (1-G)\left(x_{t+1|t} - x_{t+1|t-1}\right) + G v_{t+1}$$

$$\Rightarrow x_{t+1} - x_{t+1|t} = \frac{\sigma_G^2}{G} \left\{ \sum_{k=0}^{\infty} A_k(G) x_{t-k-1} \right\} + \frac{(1-G)}{G}\left(x_{t+1|t} - x_{t+1|t-1}\right) + v_{t+1}$$

so forecast errors should be predictable using lagged values of the variable being forecasted in addition to forecast revisions.

Appendix D: Noisy-Information Model with Heterogeneous Priors about Long-Run Means

The setup is the same as in the noisy-information model of section 2.2, but forecasters report forecasts

$$F_{it}x_{t+h} = \omega\mu_i + (1-\omega)x_{t+h|t}(i)$$

where, following Patton and Timmermann (2010), the shrinkage factor ω is common across agents, and $x_{t+h|t}(i) = E(x_{t+h}|\Omega_{it})$ is agent i's conditional expectation of x from the Kalman filter given private information Ω_{it}.

As a result, the average reported forecast across agents is given by (assuming $E_i(\mu_i) = 0$):

$$F_t x_{t+h} = (1-\omega)x_{t+h|t}$$

Averaging across Kalman forecasts of agents, we previously had:

$$x_{t+h} - x_{t+h|t} = \frac{1-G}{G}\left(x_{t+h|t} - x_{t+h|t-1}\right) + REerror_t,$$

where $REerror_t$ is the rational expectations error.

After substituting for reported forecasts, we get

$$x_{t+h} - \left(\frac{1}{1-\omega}\right)F_t x_{t+h} = \frac{1-G}{G}\left(\left(\frac{1}{1-\omega}\right)F_t x_{t+h} - \left(\frac{1}{1-\omega}\right)F_{t-1}x_{t+h}\right) + REerror_t$$

$$x_{t+h} - F_t x_{t+h} = \frac{1-G}{G}(F_t x_{t+h} - F_{t-1}x_{t+h}) + \omega x_{t+h} + REerror_t$$

$$x_{t+h} - F_t x_{t+h} = \frac{1-G}{G}(F_t x_{t+h} - F_{t-1}x_{t+h}) + \omega\rho^{h+1}x_{t-1} + v_t + REerror_t$$

Hence, the ex-post forecast error now depends on ex-ante forecast revisions as well as the lagged forecast level. In addition, the error now includes a time *t* component which will generally be correlated with current forecasts.

Appendix E: Heterogeneity in Loss Aversion

Following Capistran and Timmermann (2009), suppose that inflation follows: $\pi_t = \rho\pi_{t-1} + v_t$ where v_t is not serially correlated but potentially heteroskedastic. Specifically, $\sigma_t = \alpha_0 + \alpha_1 v_{t-1}^2 + \beta\sigma_{t-1}$ so that $v_t \sim (0, \sigma_t)$. The mean forecast is given by

$$\overline{\pi_{t+s|t}} = \rho^s \pi_t + \frac{1}{2}\sigma_{t+s|t}^2 \bar{E}(\phi_i)$$

where ϕ_i measures loss-aversion for agent *i* based on the LINEX function in section 2.4.4, $\bar{E}(\phi_i)$ is the average value of ϕ_i across agents.

Note that

$$\begin{aligned}\overline{\pi_{t+2|t}} &= \rho^2\pi_t + \tfrac{1}{2}\bar{E}(\phi_i)\sigma_{t+2|t}^2 \\ &= \rho^2\pi_t + \tfrac{1}{2}\bar{E}(\phi_i)E_t\left\{\alpha_0 + \alpha_1 v_{t+1}^2 + \beta\{\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t\}\right\} \\ &= \rho^2\pi_t + \tfrac{1}{2}\bar{E}(\phi_i)\left\{\alpha_0 + \alpha_1[\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t] + \beta\{\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t\}\right\}\end{aligned}$$

$$\begin{aligned}\overline{\pi_{t+2|t+1}} &= \rho^2\pi_t + \rho v_{t+1} + \tfrac{1}{2}\bar{E}(\phi_i)\sigma_{t+2|t+1}^2 \\ &= \rho^2\pi_t + \rho v_{t+1} + \tfrac{1}{2}\bar{E}(\phi_i)E_{t+1}\{\alpha_0 + \alpha_1 v_{t+1}^2 + \beta\sigma_{t+1}\} \\ &= \rho^2\pi_t + \rho v_{t+1} + \tfrac{1}{2}\bar{E}(\phi_i)\left\{\alpha_0 + \alpha_1 v_{t+1}^2 + \beta\{\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t\}\right\}\end{aligned}$$

Forecast error is given by

$$\begin{aligned}\pi_{t+2} - \overline{\pi_{t+2|t+1}} &= v_{t+2} - \tfrac{1}{2}\bar{E}(\phi_i)\left\{\alpha_0 + \alpha_1 v_{t+1}^2 + \beta\{\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t\}\right\} \\ &= v_{t+2} - \tfrac{1}{2}\bar{E}(\phi_i)\{\alpha_0 + \alpha_1 v_{t+1}^2 + \beta\sigma_{t+1}\}\end{aligned}$$

The revision of the forecast is given by

$$\begin{aligned}\overline{\pi_{t+2|t+1}} - \overline{\pi_{t+2|t}} &= \rho v_{t+1} + \tfrac{1}{2}\bar{E}(\phi_i)\left\{\alpha_0 + \alpha_1 v_{t+1}^2 + \beta\{\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t\}\right\} \\ &\quad -\tfrac{1}{2}\bar{E}(\phi_i)\left\{\alpha_0 + \alpha_1[\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t] + \beta\{\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t\}\right\} \\ &= \rho v_{t+1} + \tfrac{1}{2}\alpha_1\bar{E}(\phi_i)\{v_{t+1}^2 - [\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t]\} \\ &= \rho v_{t+1} + \tfrac{1}{2}\alpha_1\bar{E}(\phi_i)\{v_{t+1}^2 - \sigma_{t+1}\}\end{aligned}$$

That is revision in the forecast is a function of the shock received at time t+1. More specifically, the level of the shock (v_{t+1}) and how its volatility is different from expected volatility ($\{v_{t+1}^2 - [\alpha_0 + \alpha_1 v_t^2 + \beta\sigma_t]\}$).

Consider regressing forecast error $\pi_{t+2} - \overline{\pi_{t+2|t+1}}$ on forecast revision $\overline{\pi_{t+2|t+1}} - \overline{\pi_{t+2|t}}$.
Note that if the distribution of shocks is symmetric then $(v_{t+1}, v_{t+1}^2) = 0$. Also, by assumption, $cov(v_{t+1}, v_{t+2}) = 0$. Finally, σ_{t+1} is predetermined at time t and thus $E[v_{t+1}^2\sigma_{t+1}] = E\left(E_t(v_{t+1}^2\sigma_{t+1})\right) = E\left(\sigma_{t+1}E_t(v_{t+1}^2)\right) = E(\sigma_{t+1}\sigma_{t+1}) = E(\sigma_{t+1}^2) \equiv S$

Thus the sign of the regression coefficient is given by

$$\begin{aligned}&E\{\left[-\tfrac{1}{2}\bar{E}(\phi_i)\{\alpha_0 + \alpha_1 v_{t+1}^2 + \beta\sigma_{t+1}\}\right]\left[\tfrac{1}{2}\alpha_1\bar{E}(\phi_i)\{v_{t+1}^2 - \sigma_{t+1}\}\right]\} = \\ &= -\tfrac{1}{4}\alpha_1[\bar{E}(\phi_i)]^2 E\{[\alpha_1 v_{t+1}^2 + \beta\sigma_{t+1}][v_{t+1}^2 - \sigma_{t+1}]\} = \\ &= -\tfrac{1}{4}\alpha_1[\bar{E}(\phi_i)]^2\{\alpha_1 E(v_{t+1}^4) - \alpha_1 S\} \\ &= -\tfrac{1}{4}\alpha_1^2[\bar{E}(\phi_i)]^2 E(\{v_{t+1}^2 - \sigma_{t+1}\}^2).\end{aligned}$$

Therefore, forecast errors and forecast revisions should be *negatively* correlated.

Appendix F: Dynamic Forecast Smoothing

We want to consider the optimization problem of a forecaster who seeks to minimize the mean squared forecast error but also wishes to minimize changes in the forecasts for reputational considerations. At time *t*, given a forecast from time $t-1$, the forecaster needs to choose a sequence of forecasts (in expectation) of outcomes of a variable *x* at time $t+h$

$$min \sum_{j=0}^{h} \gamma^j E_t \left[\left(x_{t+h} - F_{t+j}x_{t+h} \right)^2 + \alpha \left(F_{t+j}x_{t+h} - F_{t+j-1}x_{t+h} \right)^2 \right],$$

where γ is the discount factor and the problem is bounded in time by $t+h$, since in subsequent periods the actual data will be revealed.

The first-order condition (FOC) is

$$E_t x_{t+h} - E_t F_{t+j} x_{t+h} = \alpha \left(E_t F_{t+j} x_{t+h} - E_t F_{t+j-1} x_{t+h} \right) - \alpha\gamma \left(E_t F_{t+j+1} x_{t+h} - E_t F_{t+j} x_{t+h} \right)$$

for all $j < h$. At horizon $j = h$, the FOC is the same as in the static model.

Consider the case of $j = 0$, i.e. choosing current forecast, we have

$$E_t x_{t+h} - F_t x_{t+h} = \alpha (F_t x_{t+h} - F_{t-1} x_{t+h}) - \alpha\gamma (E_t F_{t+1} x_{t+h} - F_t x_{t+h})$$

such that if we impose full-information rational expectations, we get

$$x_{t+h} - F_t x_{t+h} = \alpha (F_t x_{t+h} - F_{t-1} x_{t+h}) - \alpha\gamma (F_{t+1} x_{t+h} - F_t x_{t+h}) + REerror_t$$

$$\Leftrightarrow x_{t+h} - F_{t+1} x_{t+h} = -(1+\alpha\gamma)(F_{t+1} x_{t+h} - F_t x_{t+h}) + \alpha (F_t x_{t+h} - F_{t-1} x_{t+h}) + REerror_t$$

or using the same timing as for other specifications in the text

$$x_{t+h} - F_t x_{t+h} = -(1+\alpha\gamma)(F_t x_{t+h} - F_{t-1} x_{t+h}) + \alpha (F_{t-1} x_{t+h} - F_{t-2} x_{t+h}) + REerror_{t-1}$$

where $REerror_{t-1}$ is the rational expectations error such that $E[REerror_{t-1} Z_{t-1}] = 0$, i.e. the error is orthogonal to information dated $t-1$ and earlier but not time t.

With $\alpha > 0, \beta > 0$, dynamic forecast smoothing implies a negative coefficient on contemporaneous forecast revision and positive on lagged forecast revision. The error term will in general be correlated with information dated time *t* so OLS is inapplicable.

Appendix G: Aggregation

Suppose information rigidity varies across some dimension (across variables, horizons, time periods). In this appendix, we derive a relationship between information rigidity estimated for each case (variable/horizon/period) separately and information rigidity estimated on a sample pooling across some dimension or dimensions.

To make intuition transparent, suppose there are two sub-samples where, without loss of generality, information rigidity varies over time and each subsample is characterized by a different degree of information rigidity. For each subsample, the sensitivity of forecast errors (FE) to forecast revisions (FR) is described by

Subsample #1: $FE_t = \beta_1 FR_t + e_t^{(1)}$

Subsample #2: $FE_t = \beta_2 FR_t + e_t^{(2)}$

where β_1 and β_2 are measures of information rigidity in each subsample, $e_t^{(1)}$ and $e_t^{(2)}$ are rational expectations errors. The sizes of subsamples #1 and #2 is T_1 and T_2, with $T \equiv T_1 + T_2$. One can estimate β_1 and β_2 consistently by OLS for each subsample.

When one estimates the specification on the pooled sample, the estimate is given by

$$\hat{\beta}^{pool} = \frac{\frac{1}{T}\sum_{t=1}^{T} FE_t FR_t}{\frac{1}{T}\sum_{t=1}^{T} FR_t^2} = \frac{\frac{1}{T}\frac{T_1}{T_1}\sum_{t=1}^{T_1} FE_t FR_t + \frac{1}{T}\frac{T_2}{T_2}\sum_{t=T_1+1}^{T} FE_t FR_t}{\frac{1}{T}\frac{T_1}{T_1}\sum_{t=1}^{T_1} FR_t^2 + \frac{1}{T}\frac{T_2}{T_2}\sum_{t=T_1+1}^{T} FR_t^2} =$$

$$= \frac{\frac{T_1}{T}\frac{\frac{1}{T_1}\sum_{t=1}^{T_1} FE_t FR_t}{\frac{1}{T_1}\sum_{t=1}^{T_1} FR_t^2}\left(\frac{1}{T_1}\sum_{t=1}^{T_1} FR_t^2\right) + \frac{T_2}{T}\frac{\frac{1}{T_2}\sum_{t=T_1+1}^{T} FE_t FR_t}{\frac{1}{T_2}\sum_{t=T_1+1}^{T} FR_t^2}\left(\frac{1}{T_2}\sum_{t=T_1+1}^{T} FR_t^2\right)}{\frac{1}{T}\frac{T_1}{T_1}\sum_{t=1}^{T_1} FR_t^2 + \frac{1}{T}\frac{T_2}{T_2}\sum_{t=T_1+1}^{T} FR_t^2}$$

$$= \frac{\frac{T_1}{T}\hat{\beta}_1\left(\frac{1}{T_1}\sum_{t=1}^{T_1} FR_t^2\right) + \frac{T_2}{T}\hat{\beta}_2\left(\frac{1}{T_2}\sum_{t=T_1+1}^{T} FR_t^2\right)}{\frac{1}{T}\frac{T_1}{T_1}\sum_{t=1}^{T_1} FR_t^2 + \frac{1}{T}\frac{T_2}{T_2}\sum_{t=T_1+1}^{T} FR_t^2}$$

$$= \frac{\frac{T_1}{T}\hat{\beta}_1 \mathrm{var}_{FR}^{(1)} + \frac{T_2}{T}\hat{\beta}_2 \mathrm{var}_{FR}^{(2)}}{\frac{T_1}{T}\mathrm{var}_{FR}^{(1)} + \frac{T_2}{T}\mathrm{var}_{FR}^{(2)}}$$

where $\mathrm{var}_{FR}^{(1)} \equiv \frac{1}{T_1}\sum_{t=1}^{T_1} FR_t^2$ and $\mathrm{var}_{FR}^{(2)} \equiv \frac{1}{T_2}\sum_{t=T_1+1}^{T} FR_t^2$. It is straightforward to generalize the expression for the case with N subsamples.

In short, $\hat{\beta}^{pool}$ is a weighted average of individual information rigidities where weights depend on the size of the sample and the amount of variation in the data. $\hat{\beta}^{pool}$ can be interpreted as a summary measure (average) of information rigidity aggregating potentially heterogeneous information rigidity across some dimension(s).

Appendix H: Mapping into Rational Inattention

Mackowiak and Wiederholt (2009) show that the average mean squared error (MSE) of contemporaneous forecast errors for variable X is

$$E[(X_t - Y_t^*)^2] = \sigma_X^2 \times \frac{1-\rho^2}{2^{2\kappa_j} - \rho^2}$$

when an agent chooses κ_j optimally (bits to allocate to the macro variable). For inflation, we can measure the left-hand side of the above equation directly: this is the MSE of SPF inflation forecasts. In the data (Appendix Table 1), the MSE for SPF inflation forecasts at horizon zero (to make it directly comparable to the equation in Mackowiak and Wiederholt) is 1.146^2. σ_X^2 is the unconditional variance of inflation (GDP deflator; quarter on quarter). In the data, it is equal to 2.453^2. The serial correlation of inflation (GDP deflator; quarter on quarter) is 0.891. Using these figures, we obtain that $2^{2\kappa_j}$ =1.741 and therefore the gain of the Kalman filter $G = 0.42$ and implied information rigidity is $1 - G = 0.58$. This is the value provided in footnote 7.