Technical appendix for:

"Why do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion"

by Andrew J. Patton and Allan Timmermann

27 July 2010

Technical Appendix

This appendix derives the moments used in the empirical estimation in Section 3. The state and measurement equations underlying the model from Section 3 are used to show how the forecasters' updating equations can be solved.

A.1. State and Measurement Equations

Our model involves unobserved variables and can be cast in state space form using notation similar to that in Hamilton (1994). To account for the way the target variable is constructed, $z_t \equiv \Sigma_{j=0}^{11} y_{t-j}$, the state equation is augmented with eleven lags of y_t so the target variable can be written as a linear combination of the state variable. The state equation is

$$\begin{bmatrix} x_t \\ y_t \\ y_{t-1} \\ \vdots \\ y_{t-11} \end{bmatrix} = \begin{bmatrix} \phi & 0 & 0 & \cdots & 0 \\ \phi & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-12} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \varepsilon_t + u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{1}$$

which can be written as

$$\boldsymbol{\xi}_t = \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t. \tag{2}$$

The measurement equation involves two variables: the estimate of y_t incorporating both common and idiosyncratic measurement error, and the estimate of y_{t-1} incorporating just common measurement error. In a minor abuse of notation relative to our discussion of this model in Section 3, the former is labeled $\tilde{y}^*_{it}$ and the latter is labeled $\tilde{y}_{c,t-1}$, so they can be stacked into a vector $\tilde{\mathbf{y}}_{it}$:

$$\begin{bmatrix} \tilde{y}^*_{it} \\ \tilde{y}_{c,t-1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ y_{t-1} \\ \vdots \\ y_{t-11} \end{bmatrix} + \begin{bmatrix} \eta_t + \nu_{it} \\ \varphi_{t-1} \end{bmatrix} \tag{3}$$

which is written as

$$\tilde{\mathbf{y}}_{it} = \mathbf{H}'\boldsymbol{\xi}_t + \mathbf{w}_{it},$$

To simplify the model, we introduce the measurement error φ_{t-1}, distinct from η_t but with the same distribution, so that the vector $\mathbf{w}_{it}$ remains serially uncorrelated.

The various shocks in the state and measurement equations are distributed as:

$$(u_t, \varepsilon_t, \eta_t, \varphi_t, \nu_{1t}, ..., \nu_{NT})' \sim iid\ N\left(\mathbf{0}, diag\left\{\left(\sigma_u^2, \sigma_\varepsilon^2, \sigma_\eta^2, \sigma_\eta^2, \sigma_\nu^2, ..., \sigma_\nu^2\right)\right\}\right)$$

where $diag\{\mathbf{a}\}$ is a square diagonal matrix with the vector $\mathbf{a}$ on the main diagonal. Then $\mathbf{v}_t \sim iid\ N(0, \mathbf{Q})$, with

$$\mathbf{Q} = \begin{bmatrix} \sigma_\varepsilon^2 & \sigma_\varepsilon^2 & 0 & \cdots & 0 \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 + \sigma_u^2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

And finally $\mathbf{w}_{it} \sim iid\ N(0, \mathbf{R})$, with

$$\mathbf{R} = \begin{bmatrix} \sigma_\eta^2 + \sigma_\nu^2 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix}.$$

Notice that by extending the state variable to include lags of $\mathbf{y}_t$, forecasts, nowcasts and backcasts need not be treated separately. They can all be treated simultaneously as "forecasts" of the state vector $\boldsymbol{\xi}_t$. This simplifies the algebra considerably.

A.2. The Forecasters' Updating Process

Our empirical data provide us with estimates of forecast uncertainty at different forecast horizons measured both in the form of the root mean squared forecast error (RMSE) of the "average" or consensus forecast or in the form of the cross-sectional standard deviation of the forecasts (i.e., the dispersion). The analysis next characterizes how the forecasters update their beliefs and derives the model-implied counterparts of these two measures of uncertainty and disagreement.

Let

$$\begin{aligned} \mathcal{F}_{it} &= \sigma\left(\tilde{\mathbf{y}}_{it}, \tilde{\mathbf{y}}_{i,t-1}, ..., \tilde{\mathbf{y}}_{i1}\right) \\ \hat{\boldsymbol{\xi}}_{it|t-h} &\equiv E\left[\boldsymbol{\xi}_t | \mathcal{F}_{i|t-h}\right], \quad h \geq 0, \end{aligned}$$

where the expectation is obtained using standard Kalman filtering methods.

Forecasters are assumed to have used the Kalman filter long enough that all updating matrices, defined below, are at their steady-state values. This is done simply to remove any start of sample effects that could be present in the data. Following Hamilton (1994):

$$\begin{aligned}
\mathbf{P}_{i,t+1|t} &\equiv E\left[\left(\boldsymbol{\xi}_{t+1}-\hat{\boldsymbol{\xi}}_{i,t+1|t}\right)\left(\boldsymbol{\xi}_{t+1}-\hat{\boldsymbol{\xi}}_{i,t+1|t}\right)'\right] \\
&= \left(\mathbf{F}-\mathbf{K}_{it}\right)\mathbf{P}_{it|t-1}\left(\mathbf{F}'-\mathbf{K}'_{i,t}\right)+\mathbf{K}_{it}\mathbf{R}\mathbf{K}'_{it}+\mathbf{Q} \\
&\to \mathbf{P}_1^*. \qquad (4)
\end{aligned}$$

Note that although the individual forecasters receive different signals, and thus generate different forecasts $\hat{\boldsymbol{\xi}}_{i,t+1|t}$, all signals have the same covariance structure and so will converge to the same matrix, $\mathbf{P}_1^*$. Similarly,[1]

$$\begin{aligned}
\mathbf{K}_{it} &\equiv \mathbf{F}\mathbf{P}_{it|t-1}\left(\mathbf{P}_{it|t-1}+\mathbf{R}\right)^{-1} \to \mathbf{K}^*, \\
\mathbf{P}_{it|t} &\equiv E\left[\left(\boldsymbol{\xi}_t-\hat{\boldsymbol{\xi}}_{it|t}\right)\left(\boldsymbol{\xi}_t-\hat{\boldsymbol{\xi}}_{it|t}\right)'\right] \\
&= \mathbf{P}_{it|t-1}-\mathbf{P}_{it|t-1}\left(\mathbf{P}_{it|t-1}+\mathbf{R}\right)^{-1}\mathbf{P}_{it|t-1} \\
&\to \mathbf{P}_1^*-\mathbf{P}_1^*\left(\mathbf{P}_1^*+\mathbf{R}\right)^{-1}\mathbf{P}_1^* \equiv \mathbf{P}_0^*. \qquad (5)
\end{aligned}$$

To estimate the matrices $\mathbf{P}_1^*$, $\mathbf{P}_0^*$, and $\mathbf{K}^*$, 100 non-overlapping years of data are simulated and $\mathbf{P}_{it|t-1}$, $\mathbf{P}_{it|t}$ and $\mathbf{K}_{it}$ are updated using the above equations. These matrices at the end of the 100^{th} year are used as estimates of $\mathbf{P}_1^*$, $\mathbf{P}_0^*$, and $\mathbf{K}^*$. Multi-step prediction errors use

$$\begin{aligned}
\hat{\boldsymbol{\xi}}_{i,t+h|t} &= \mathbf{F}^h\hat{\boldsymbol{\xi}}_{it|t}, \\
\text{so } \mathbf{P}_{i,t+h|t} &\equiv E\left[\left(\boldsymbol{\xi}_{t+h}-\hat{\boldsymbol{\xi}}_{i,t+h|t}\right)\left(\boldsymbol{\xi}_{t+h}-\hat{\boldsymbol{\xi}}_{i,t+h|t}\right)'\right] \qquad (6) \\
&= \mathbf{F}^h\mathbf{P}_{it|t}\left(\mathbf{F}'\right)^h+\sum_{j=0}^{h-1}\mathbf{F}^j\mathbf{Q}\left(\mathbf{F}'\right)^j \to \mathbf{P}_h^*, \text{ for } h\geq 1.
\end{aligned}$$

The matrices $\mathbf{P}_h^*$ for $h = 1, 2, ..., 24$ are sufficient to obtain the term structure of RMSE, (that is, the RMSE-values across different horizons, $h = 1, .., H$), for an individual forecaster, but the moments included in the estimation are from the consensus forecasts, and so the RMSE term structure for the consensus is needed.[2] To this end, let

$$\bar{\boldsymbol{\xi}}_{t|t-h} \equiv \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\xi}}_{it|t-h} \qquad (7)$$

[1] The convergence of $\mathbf{P}_{it|t-1}$, $\mathbf{P}_{it|t}$ and $\mathbf{K}_{it}$ to their steady-state values relies on $|\phi| < 1$, see Hamilton (1994), Proposition 13.1, and this is imposed in the estimation.

[2] Patton and Timmermann (2010) also consider the behavior of the consensus forecast error but do not analyze cross-sectional dispersion in forecasts.

be the consensus forecast of the state vector. Before deriving the term structure of RMSE for this forecast, it is useful to derive the RMSE of the consensus "nowcast":

$$\begin{aligned}
\bar{\mathbf{P}}_0^* &\equiv V\left[\boldsymbol{\xi}_t - \bar{\boldsymbol{\xi}}_{t|t}\right] \\
&= V\left[\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t}\right)\right] \\
&= \frac{1}{N^2}\sum_{i=1}^{N} V\left[\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t}\right] + \frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{k=i+1}^{N} Cov\left[\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t}, \boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{kt|t}\right] \\
&= \frac{1}{N}\mathbf{P}_0^* + \frac{N-1}{N} E\left[\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t}\right)\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{kt|t}\right)'\right], \qquad (8)
\end{aligned}$$

using the assumption that all forecasters receive signals with identical distributions. It is possible to show that the current nowcast error is the following function of the previous period's nowcast error and the intervening innovations:

$$\begin{aligned}
\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t} &= \left(\mathbf{I} - \mathbf{P}_1^*\mathbf{H}\left(\mathbf{H}'\mathbf{P}_1^*\mathbf{H} + \mathbf{R}\right)^{-1}\mathbf{H}'\right)\mathbf{F}\left(\boldsymbol{\xi}_{t-1} - \hat{\boldsymbol{\xi}}_{i,t-1|t-1}\right) \\
&\quad + \left(\mathbf{I} - \mathbf{P}_1^*\mathbf{H}\left(\mathbf{H}'\mathbf{P}_1^*\mathbf{H} + \mathbf{R}\right)^{-1}\mathbf{H}'\right)\mathbf{v}_t \\
&\quad - \mathbf{P}_1^*\mathbf{H}\left(\mathbf{H}'\mathbf{P}_1^*\mathbf{H} + \mathbf{R}\right)^{-1}\mathbf{w}_{it} \\
&\equiv \mathbf{A}\left(\boldsymbol{\xi}_{t-1} - \hat{\boldsymbol{\xi}}_{i,t-1|t-1}\right) + \mathbf{B}\mathbf{v}_t + \mathbf{C}\mathbf{w}_{it}, \qquad (9)
\end{aligned}$$

where $\mathbf{v}_t$ and $\mathbf{w}_{it}$ are defined above. This result is used to derive the covariance between nowcast errors across different forecasters:

$$\begin{aligned}
\mathbf{P}_{0ik}^* &\equiv E\left[\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t}\right)\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{kt|t}\right)'\right] \qquad (10) \\
&= E\left[\left(\mathbf{A}\left(\boldsymbol{\xi}_{t-1} - \hat{\boldsymbol{\xi}}_{i,t-1|t-1}\right) + \mathbf{B}\mathbf{v}_t + \mathbf{C}\mathbf{w}_{it}\right)\left(\mathbf{A}\left(\boldsymbol{\xi}_{t-1} - \hat{\boldsymbol{\xi}}_{k,t-1|t-1}\right) + \mathbf{B}\mathbf{v}_t + \mathbf{C}\mathbf{w}_{kt}\right)'\right] \\
&= \mathbf{A}E\left[\left(\boldsymbol{\xi}_{t-1} - \hat{\boldsymbol{\xi}}_{i,t-1|t-1}\right)\left(\boldsymbol{\xi}_{t-1} - \hat{\boldsymbol{\xi}}_{k,t-1|t-1}\right)'\right]\mathbf{A}' + \mathbf{B}\mathbf{Q}\mathbf{B}' + \mathbf{C}E\left[\mathbf{w}_{it}\mathbf{w}_{kt}'\right]\mathbf{C}',
\end{aligned}$$

with all other terms in the two nowcast errors having zero covariance. Letting

$$E\left[\mathbf{w}_{it}\mathbf{w}_{kt}'\right] = \begin{bmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix} \equiv \mathbf{R}_{ik},$$

it follows that

$$\mathbf{P}_{0ik}^* = \mathbf{A}\mathbf{P}_{0ik}^*\mathbf{A}' + \mathbf{B}\mathbf{Q}\mathbf{B}' + \mathbf{C}\mathbf{R}_{ik}\mathbf{C}',$$

which exploits the stationarity of the process, and yields an implicit solution for the covariance of nowcast errors across forecasters, $\mathbf{P}^*_{0ik}$.[3] Thus the variance of the error of the consensus nowcast of the state vector is:

$$\bar{\mathbf{P}}^*_0 \equiv V\left[\boldsymbol{\xi}_t - \bar{\boldsymbol{\xi}}_{t|t}\right] = \frac{1}{N}\mathbf{P}^*_0 + \frac{N-1}{N}\mathbf{P}^*_{0ik}. \tag{11}$$

The variance of the consensus forecast of the state vector for $h \geq 1$ can be similarly obtained. Using the following expression for forecast errors as a function of a previous nowcast error and the intervening innovations

$$\boldsymbol{\xi}_t - \boldsymbol{\xi}_{it|t-h} = \mathbf{F}^h\left(\boldsymbol{\xi}_{t-h} - \boldsymbol{\xi}_{i,t-h|t-h}\right) + \sum_{j=0}^{h-1}\mathbf{F}^j\mathbf{v}_{t-j},\ h \geq 1, \tag{12}$$

it follows that

$$\begin{aligned}
\bar{\mathbf{P}}^*_h &\equiv V\left[\boldsymbol{\xi}_t - \bar{\boldsymbol{\xi}}_{t|t-h}\right] \\
&= V\left[\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t-h}\right)\right] \\
&= \frac{1}{N^2}\sum_{i=1}^{N} V\left[\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t-h}\right] + \frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{k=i+1}^{N} Cov\left[\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t-h}, \boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{kt|t-h}\right] \\
&= \frac{1}{N}\mathbf{P}^*_h + \frac{N-1}{N}E\left[\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t-h}\right)\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{kt|t-h}\right)'\right] \qquad (13)
\end{aligned}$$

To evaluate this expression requires knowledge of the covariance between the individual forecast errors measured at different horizons:

$$\begin{aligned}
\mathbf{P}^*_{hik} &\equiv E\left[\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{it|t-h}\right)\left(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{kt|t-h}\right)'\right] \\
&= E\left[\left(\mathbf{F}^h\left(\boldsymbol{\xi}_{t-h} - \boldsymbol{\xi}_{i,t-h|t-h}\right) + \sum_{j=0}^{h-1}\mathbf{F}^j\mathbf{v}_{t-j}\right)\left(\mathbf{F}^h\left(\boldsymbol{\xi}_{t-h} - \boldsymbol{\xi}_{k,t-h|t-h}\right) + \sum_{j=0}^{h-1}\mathbf{F}^j\mathbf{v}_{t-j}\right)'\right] \\
&= \mathbf{F}^h E\left[\left(\boldsymbol{\xi}_{t-h} - \boldsymbol{\xi}_{i,t-h|t-h}\right)\left(\boldsymbol{\xi}_{t-h} - \boldsymbol{\xi}_{k,t-h|t-h}\right)'\right]\left(\mathbf{F}^h\right)' \\
&\quad + E\left[\left(\sum_{j=0}^{h-1}\mathbf{F}^j\mathbf{v}_{t-j}\right)\left(\sum_{j=0}^{h-1}\mathbf{F}^j\mathbf{v}_{t-j}\right)'\right] \qquad (14) \\
&= \mathbf{F}^h\mathbf{P}^*_{0ik}\left(\mathbf{F}^h\right)' + \sum_{j=0}^{h-1}\mathbf{F}^j\mathbf{Q}\left(\mathbf{F}^j\right)',\ \ h \geq 1. \qquad (15)
\end{aligned}$$

[3] Like other covariance matrices that appear in more standard Kalman filtering applications, see Hamilton (1994), Proposition 13.1 for example, it is not possible to obtain an explicit expression for $\mathbf{P}^*_{0ik}$.

With these moment matrices in place it is simple to obtain the term structure of MSE-values for the consensus forecast of the target variable. Let $\boldsymbol{\omega} \equiv [0, \boldsymbol{\iota}_{12}']'$, where $\boldsymbol{\iota}_k$ is a $k \times 1$ vector of ones, so:

$$V\left[z_t - \bar{z}_{t|t-h}\right] = V\left[\boldsymbol{\omega}'\left(\boldsymbol{\xi}_t - \bar{\boldsymbol{\xi}}_{t|t-h}\right)\right] = \boldsymbol{\omega}'\bar{\mathbf{P}}_h^*\boldsymbol{\omega}, \text{ for } h \geq 0. \tag{16}$$

The above expression yields 24 moments (the mean squared errors for the 24 forecast horizons) that can be used to estimate the parameters of the model governing the dynamics of GDP growth and inflation.

References

Hamilton, J.D., 1994, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.

Patton, A.J. and A. Timmermann, 2010, Predictability of Output Growth and Inflation: A Multi-horizon Survey Approach. Forthcoming in *Journal of Business and Economic Statistics*.