Dynamic Hierarchical Factor Models*

Emanuel Moench† Serena Ng‡ Simon Potter§

April 14, 2010

Abstract

This paper uses multi-level factor models to characterize within and between block variations as well as idiosyncratic noise in large dynamic panels. Block-level shocks are distinguished from genuinely common shocks, and the estimated block-level factors are easy to interpret. The framework achieves dimension reduction and yet explicitly allows for heterogeneity between blocks. The model is estimated using a MCMC algorithm that takes into account the hierarchical structure of the factors. We organize a panel of 447 series into blocks according to the timing of data releases and use a four level model to study the dynamics of real activity at both the block and aggregate levels. While the economic downturn of 2007-2009 is pervasive, growth cycles are only loosely synchronized across blocks. The state of the leading and the lagging sectors, as well as the overall economy are monitored in a coherent framework.

Keywords: forecasting, monitoring, co-movements, large dimensional panel, diffusion index.

JEL classification: C10, C20, C30

*We would like to thank Evan LeFlore and Jenny Chan for excellent research assistance on this project. The second author would like to acknowledge financial support from the National Science Foundation under grant SES 0549978. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System.

†Federal Reserve Bank of New York, 33 Liberty St., New York, NY 10045. emanuel.moench@ny.frb.org

‡Columbia University, 420 W. 118 St., MC 3308, New York, NY 10025. serena.ng@columbia.edu

§Federal Reserve Bank of New York, 33 Liberty St., New York, NY 10045. simon.potter@ny.frb.org

1 Introduction

Recent research has found that in a data rich environment, dimension reduction in the form of factors is useful for forecasting and policy analysis. However, a common criticism of factor models is that the factors are difficult to interpret. One reason is that the factors are estimated from a large panel of data without taking full advantage of the data structure. This paper proposes a new hierarchical (multi-level) dynamic factor model obtained by splitting a large panel of data into a small number of blocks. In our three level model, the within-block comovements are due to block-level factors, and the between-block variations are due to the common factors. Factors estimated from a block of employment data, for example, are then interpreted as a labor market factor. A four level model further divides the blocks into subblocks. The hierarchical structure implies that the transition equation for the factors at a given level has a time varying intercept that depends on the factors at the next higher level. We show how this can be taken into account in state space estimation.

We consider real time monitoring of economic activity as a primary use of our model. Timely and accurate monitoring of the data are necessary for policy makers to design their policies in accordance with changing economic conditions. At least three issues make the monitoring exercise difficult in practice, however. First, as information technology changes, more and more economic data become available. Second, as individual time series typically feature idiosyncratic variation and are measured with error, they contain not only news about the state of the economy but also noise. Third, the data are released at different points in time. Real time monitoring of economic activity therefore requires filtering out the common signal from the idiosyncratic noise for a very large number of variables in inherently unbalanced panels. By exploiting the timing in which the data are released, our model provides a formal way to aggregate information at different levels as blocks of data arrive on a staggered basis. We use a four level model to monitor 447 economic time series and find that the timing of the 2007-2009 downturn and subsequent recovery varies quite substantially across different blocks. Moreover, a continuous update of the state of the economy based on each new release of data provides a more timely signal of growth dynamics than an update at the end of each month when new observations on all blocks of data have been released.

Our model can be applied whenever a panel of data can sensibly be organized into blocks and block level variations are non-negligible. Indeed, the decomposition of variances show that block-level variations tend to be stronger than the common variations, though both are small relative to the purely idiosyncratic components in most series. We also compare the factor estimates to the principal component estimates and show that the principal components estimator tends to treat block-level variation as common because the large magnitude of the block-level event contributes significantly to the total variations in the data. Our model can also be used to analyze the relative importance of country-level and global variations such as in Glick and Rogoff (1995), or market-wide versus industry-specific variations such as in Brooks and DelNegro (2006).

The remainder of this paper is organized as follows. In Section 2, we introduce the three level specification of our model. We present its hierarchical state space representation in Section 2.1, outline its estimation via Markov Chain Monte Carlo methods in Section 2.2 and discuss how our model relates to previously suggested multi-level factor models in Section 2.3. In Section 3, we present results for a three level model of real activity in the US that comprises six blocks of data organized by the timing of data releases. In Section 4, we then present a four level extension of our model which allows for common factors at the subblock level, the block level, and the aggregate level. We apply this model to a dataset which combines several hundred economic time series for real activity in the US in five blocks and fourteen subblocks of data. Each of the subblocks corresponds to a statistical release. We discuss how our model can be used to monitor economic activity in real time in Section 5. Section 6 concludes.

2 A Three Level Hierarchical Dynamic Factor Model

We assume that the data used in the analysis (denoted X_{bit}) are stationary, mean-zero, standardized to have a unit variance after possible logarithmic transformation and detrending. We assume that there are K_{Gb} common factors G_b in each block $b = 1, \ldots, B$. Let the mean zero block-level factors be G_{bjt} for $j = 1, \ldots K_{Gb}$. Hence, there is a total number $K_G = (K_{G1} + \ldots + K_{GB})$ of block-level factors. We assume that these K_G block-level factors share a total of K_F common factors F_t. Let N_b denote the number of variables in block b. This implies a total number $N = (N_1 + \ldots + N_B)$ of variables in the analysis. We assume that N and T are both large, but that B is much smaller than N.

Each time series i in a given block b is decomposed into a serially correlated idiosyncratic component, e_{Xbit}, and a common component $\lambda^i_{G.b}(L)G_{bt}$ which it shares with other variables in the same block. Each block-level factor G_{bjt} has a serially correlated block-specific component e_{Gbjt} and a common component $\lambda^j_{F.b}(L)F_t$ which it shares with all other blocks. Finally, the economy-wide factors F_t are assumed to be serially correlated. The model can be summarized by the following equations:

$$X_{bit} = \lambda^i_{G.b}(L)G_{bt} + e_{Xbit}, \quad (1)$$

$$G_{bjt} = \lambda^j_{F.b}(L)F_t + e_{Gbjt}, \quad (2)$$

$$\psi_{F.k}(L)F_{kt} = \epsilon_{Fkt}, \quad (3)$$

$$\psi_{X.bi}(L)e_{Xbit} = \epsilon_{Xbit}, \quad (4)$$

$$\psi_{G.bj}(L)e_{Gbjt} = \epsilon_{Gbjt}. \quad (5)$$

In the above, $i = 1, \ldots N_b$, $j = 1, \ldots K_{Gb}$, and $k = 1, \ldots K_F$. In this model, variables within a block can be correlated through F_t or the e_{Gbjt}'s, but variables between blocks can be correlated only through F_t. The factor loadings and AR processes for the common factors, block-specific and idiosyncratic components are defined in more detail in Section 2.1 below. Our model specifies a multi-level structure for the factors. This is to be contrasted with models that allow for hierarchy in the factor loadings as e.g. considered in Lopes and West (2004). Such models allow for heterogeneity in the response to common variations, whereas we allow for heterogeneity in the common shocks. Related work in the literature will be further discussed in Section 2.3.

In the terminology of multilevel models, (1) is the level-one equation, and (2) is the level-two equation. The stochastic process for F_t as given by (3) constitutes a level-three equation. To fix ideas, if we are given data for production, employment, consumption, etc, then X_{1it} would be one of the N_1 series collected for production, X_{2it} would be one of the N_2 series collected for employment and so forth. The production and the employment factors G_{1t} and G_{2t} would be correlated because of economy wide fluctuations, which are captured by F_t.

The idiosyncratic components e_{Xbi} are AR processes of order q_{Xbi}, the block-specific components are AR processes of order $q_{G_{bj}}$, and the economy-wide factors F_k are AR processes of order q_{F_k}. We assume normally distributed innovations throughout. Thus,

$$\epsilon_{Xbi} \sim N(0, \sigma^2_{Xbi}) \quad i = 1, \ldots, N_b$$

$$\epsilon_{Gbj} \sim N(0, \sigma^2_{Gbj}) \quad j = 1, \ldots, K_{Gb}, \quad b = 1, \ldots, B$$

$$\epsilon_{F_k} \sim N(0, \sigma^2_{F_k}) \quad k = 1, \ldots, K_F.$$

As written, the specification allows the lag order of the factor loading matrix, the number of factors, and the factor specific errors to differ across blocks as well as within blocks. Similarly, the lag order of the idiosyncratic errors can also vary across blocks and units. The dynamics of the model could be further enriched by allowing for stochastic volatility and Markov switching effects at different levels of the hierarchy. Note, however, that when B is large, the number of possible model configurations increase quickly. In that case, restricting the parameters to be the same across blocks might be desirable.

For identification, we assume that for each $b = 1 \dots B$, the matrix of contemporaneous factor loadings $\Lambda_{G.b0} = \left(\lambda^1_{G.b0}, \lambda^2_{G.b0}, \dots, \lambda^{Nb}_{G.b0}\right)$ is lower block triangular. For example, if $K_{Gb} = 2$, we would have

$$\Lambda_{G.b0} = \begin{bmatrix} 1 & 0 \\ \lambda_{G.b0_{2,1}} & 1 \\ \lambda_{G.b0_{3,1}} & \lambda_{G.b0_{3,2}} \\ \vdots & \vdots \\ \lambda_{G.b0_{Nb,1}} & \lambda_{G.b0_{Nb,2}} \end{bmatrix}.$$

This normalization implies that the block-level factors load heavily on the variables ordered first within each block. An alternative is to normalize the variance of $\epsilon_{G.b}$ to unity and to restrict the diagonal elements of the upper-left $N_b \times K_{Gb}$ block of $\Lambda_{G.b0}$ to be positive.

Identification of the economy-wide factors F can be achieved in a similar manner. However, in the empirical examples discussed below, we assume $K_F = 1$. In this setup, it is sufficient to fix the upper-left element of the contemporaneous factor loading matrix Λ_{F0} to one in order to identify the scale of the common factor and its loadings separately.

The difference between a multilevel and a two level factor model is best understood when F_t and G_{bt} are scalars and when the factor loading polynomials $\lambda^i_{G.b}(L)$ and $\lambda^j_{F.b}(L)$ are of lag order one. With $K_{Gb} = K_F = 1$,

$$\begin{aligned} X_{bit} &= \lambda^i_{G.b}(\lambda^j_{F.b}F_t + e_{Gbt}) + e_{Xbit} \\ &= \lambda_{bi}F_t + v_{bit}, \end{aligned} \tag{6}$$

where $\lambda_{bi} = \lambda^i_{G.b}\lambda^j_{F.b}$ and

$$v_{bit} = \lambda^i_{G.b}e_{Gbt} + e_{Xbit}.$$

A standard factor model ignores the block structure and stacks all observations up irrespective of which block an observation belongs to. The data would be modeled as

$$X_{it} = \lambda_i F_t + v_{it}.$$

This two level representation would be equivalent to an exact factor model if the block-specific components $\{e_{Gbt} : b = 1, ..., B\}$ were zero for all t. We would otherwise obtain an 'approximate factor model' if v_{it} was 'weakly correlated' across i and t. In practice, this means that the number of idiosyncratic errors that are serially and/or cross-sectionally correlated cannot be too large. The condition will be satisfied if the number of series in each block was relatively small in the sense that the variation in v_{bit} is dominated by e_{Xbit} as $N \to \infty$ and $N_b \to \infty$. Instead of imposing this possibly invalid assumption, our hierarchical model tackles this problem by explicitly specifying the block structure. Since the blocks have a well-defined interpretation, and since the ordering of the first variables in each block is chosen explicitly based on knowledge about the data structure, we can provide interpretation to the block level factors. This overcomes a common criticism of large dimensional factor analysis that the factors are difficult to interpret.[1] Furthermore, estimates of G_{bt} are often objects of independent interest. For example, in monitoring the macroeconomy in our empirical analysis below, knowledge about the state of the housing block is always useful, even if the overall state of the economy might be the ultimate object of interest.

2.1 The State Space Representation

Let $\Theta = (\Theta_F; \Theta_G; \Theta_X)$ where Θ_F, Θ_G are parameters that characterize F_t, G_t respectively, and Θ_X are the remaining parameters. Let

$$
\begin{aligned}
X_{bt} &= \begin{pmatrix} X_{b.1t} & X_{b.2t} & \dots X_{b.N_b t} \end{pmatrix}' \\
X_t &= \begin{pmatrix} X_{1t} & X_{2t} & \dots X_{Bt} \end{pmatrix}' \\
G_{bt} &= \begin{pmatrix} G_{b.1t} & G_{b.2t} & \dots G_{b.K_{Gb} t} \end{pmatrix}' \\
G_t &= \begin{pmatrix} G_{1t} & G_{2t} & \dots G_{Bt} \end{pmatrix}.
\end{aligned}
$$

By assumption,

$$
(i)\ X_t \perp\!\!\!\perp \Theta | G_t, \Theta_X \quad (ii)\ G_t \perp\!\!\!\perp \Theta | F_t, \Theta_G, \quad (iii)\ F_t \perp\!\!\!\perp \Theta | \Theta_F
$$

where $\perp\!\!\!\perp$ stands for stochastic independence. Stepwise specification of the sub-models leads to the statistical model

$$
f(X_t, F_t, G_t; \Theta) = f(X_t | G_t; \Theta_X) f(G_t | F_t; \Theta_G) f(F_t; \Theta_F).
$$

The data density is

$$
f(X_t; \Theta) = \int \int f(X_t | G_t; \Theta_X) f(G_t | F_t; \Theta_G) f(F_t | \Theta_F) dG_t dF_t.
$$

[1] The widely used macroeconomic data provided by Stock and Watson (2006) is already loosely organized around blocks of data on output, consumption, prices, etc. Although it is not always clear which block some series belong to, this ambiguity does not matter as the block structure is not exploited in the analysis. In contrast, in our application data are placed into blocks by data release. For example, in the four level specification of our model we will have a retail sales block consisting of the underlying detail of the the Census Bureau's monthly retail sales release.

Because of the assumed hierarchical structure, the data density can be constructed recursively from the pair of equations:

$$\begin{aligned} f(G_t|\Theta_F,\Theta_G) &= \int f(G_t|F_t;\Theta_G)f(F_t|\Theta_F)dF_t \\ f(X_t|\Theta) &= \int f(X_t|G_t;\Theta_X)f(G_t|\Theta_F,\Theta_G)dG_t. \end{aligned}$$

Here, $f(X_t, F_t, G_t; \Theta)$ is the measurement equation and $f(G_t|\Theta_F, \Theta_G)$ is the structural model for the latent factor F_t. As discussed in Mouchart and Martin (2003), strong identification of the measurement model is required to obtain weak identification of the statistical model. Our assumptions ensure that $\Theta_X = (\Psi_X, \Sigma_X, \Lambda_G)$ are identified from the measurement model, $\Theta_G = (\Psi_G, \Sigma_G, \Lambda_F)$ are identified from the structural model for G_t, and $\Theta_F = (\Psi_F, \Sigma_F)$ are identified from the transition equation for F_t. These equations are now made precise.

Common Factor Dynamics The common factors evolve according to

$$\Psi_F(L)F_t = \epsilon_{Ft},$$

where $\Psi_F(L)$ is a $K_F \times K_F$ diagonal matrix polynomial with elements $\psi_{F.k}(L) = 1 - \psi_{F.k1}L - \ldots - \psi_{F.kq_F}L^{q_F}$ and where $\epsilon_{Ft} \sim N(0, \Sigma_F)$ and $||\vec{\Psi}_F|| < 1$.

Block-Level Dynamics The (pseudo) measurement equation that relates the block level to the common factors is

$$\begin{aligned} G_{bt} &= \Lambda_{F.b}(L)F_t + e_{Gbt}, && (7) \\ \Psi_{G.b}(L)e_{Gbt} &= \epsilon_{Gbt}. && (8) \end{aligned}$$

where $\Psi_{G.b}(L)$ is a $K_{Gb} \times K_{Gb}$ diagonal matrix polynomial with elements $\psi_{G.bj}(L) = 1 - \psi_{G.bj1}L - \ldots - \psi_{G.bjq_G}L^{q_G}$. $\Lambda_{Fb}(L)$ is a $K_{Gb} \times K_F$ matrix polynomial of factor loadings. We call this a pseudo and not a standard measurement equation because G_{bt} is not observed. We restrict $||\psi_{G.bj}(L)|| < 1$ for stationarity and assume $\epsilon_{Gt} \sim N(0, \Sigma_G)$.

Together, (7) and (8) imply that

$$\Psi_{G.b}(L)G_{bt} = \Psi_{G.b}(L)\Lambda_{F.b}(L)F_t + \epsilon_{Gbt}.$$

This leads to the block-level transition equation

$$G_{bt} = \alpha_{F.bt} + \Psi_{G.b1}G_{bt-1} + \ldots + \Psi_{G.bq_{Gb}}G_{bt-q_{Gb}} + \epsilon_{Gbt}.$$

where

$$\alpha_{F.bt} = \Psi_{G.b}(L)\Lambda_{F.b}(L)F_t. \tag{9}$$

The transition equation at the block level thus differs from the standard transition equation in linear state space models by the time-varying intercept $\alpha_{F.bt}$. This term captures the part of the dynamics of the block level factor G_{bt} that it shares with other blocks. Knowledge of the comovement across blocks is therefore useful in estimating the block-specific dynamics. In Section 2.2 below, we show how this additional term can easily be incorporated into a standard sampling method for linear state space models.

Within-Block Dynamics For each $b = 1, \ldots, B$ we have

$$\begin{aligned} X_{bt} &= \Lambda_{G.b}(L)G_{bt} + e_{Xbt}, \\ \Psi_{X.b}(L)e_{Xbt} &= \epsilon_{Xbt}, \end{aligned}$$

where $\Psi_{X.b}(L)$ is a $N_b \times N_b$ diagonal matrix polynomial with elements $\psi_{X.bi}(L) = 1 - \psi_{X.bi1}L - \ldots - \psi_{X.biq_X}L^{q_X}$. $\Lambda_{Gb}(L)$ is a $N_b \times K_{Gb}$ matrix polynomial of factor loadings with rows $\lambda^i_{G.b}(L)$. Then, for $b = 1, \ldots B$, the measurement equation for each block can be rewritten as

$$\Psi_{X.b}(L)X_{bt} = \Psi_{X.b}(L)\Lambda_{G.b}(L)G_{bt} + \epsilon_{Xbt}. \tag{10}$$

The total unconditional variance is the sum of the unconditional variance of the components multiplied by the effective loadings on the components. Dividing the variance of the components by the total variance gives the fraction of the variance in X explained by the common innovations ϵ_F, block-specific innovations ϵ_{Gb}, and idiosyncratic errors ϵ_X, respectively. We denote these by share_F, share_G, and share_X. Appendix A provides the algebraic expressions for the shares of variance due to the three types of shocks. A two level factor model does not distinguish between F_t and G_t. In these models, one minus share_X is the size of the common component.

2.2 Estimation via Markov Chain Monte Carlo

We use the method of Markov Chain Monte Carlo (MCMC) to estimate the posterior distribution of the parameters of interest. The method samples a Markov chain that has the posterior density of the parameters as its stationary distribution. MCMC has been used by Kim and Nelson (2000), Aguilar and West (2000), Geweke and Zhou (1996) and Lopes and West (2004), among others, to estimate two level factor models. These algorithms are variations and extensions of the method developed in Carter and Kohn (1994) and Frühwirth-Schnatter (1994). Although in theory, the algorithm allows for multiple factors, most previous studies have limited attention to estimation of a single factor. We allow both F_t and G_{bt} to be vector valued.

Our setup is a hierarchical dynamic factor model where each level admits a state-space representation that has a measurement and a transition equation. The MCMC algorithm thus needs to be extended to handle this hierarchical structure. Let $\mathbf{\Lambda} = (\Lambda_G, \Lambda_F)$, $\mathbf{\Psi} = (\Psi_F, \Psi_G, \Psi_X)$, $\mathbf{\Sigma} = (\Sigma_F, \Sigma_G, \Sigma_X)$. The main steps are as follows:

1. Organize the data into blocks to yield $X_{bt}, b = 1, \ldots B$. Use principal components to initialize $\{G_t\}$ and $\{F_t\}$. Use these to produce initial values for $\mathbf{\Lambda}$, $\mathbf{\Psi}$, $\mathbf{\Sigma}$.

2. Conditional on $\mathbf{\Lambda}, \mathbf{\Psi}$, $\mathbf{\Sigma}$ and $\{F_t\}$, draw $\{G_t\}$ taking into account time varying intercepts.

3. Conditional on $\mathbf{\Lambda}, \mathbf{\Psi}$, $\mathbf{\Sigma}$ and $\{G_t\}$, draw $\{F_t\}$.

4. Conditional on $\{F_t\}$ and $\{G_t\}$, draw $\mathbf{\Lambda}$, $\mathbf{\Psi}$, and $\mathbf{\Sigma}$.

5. Return to 2.

We assume conjugate priors, and thus Step (4) is straightforward. Step (3) follows the Carter and Kohn procedure used for level two models and is thus also standard. The main complication going from a two to a three level model lies in the way $\{G_t\}$ is sampled in Step (2). Recall that the transition equation for G_{bt} is of the form

$$G_{bt} = \alpha_{F.bt} + \Psi_{G.b1} G_{bt-1} + \ldots + \Psi_{G.bq_{Gb}} G_{bt-q_{Gb}} + \epsilon_{Gbt}.$$

This involves the term $\alpha_{F.bt} = \Psi_{G.b}(L)\Lambda_F(L)F_t$, which, given a draw of F_t, can be interpreted as a time-varying intercept that is known for all t. By conditioning on F_t, our updating and smoothing equations for G_t explicitly take into account the information carried by F_t.

A sketch of the algorithm is as follows. Let $\vec{G_{bt}} = \left\{G_{bt}, G_{bt-1}, \ldots, G_{bt-l_G^*}\right\}$ and define $\vec{\alpha}_{F.bt}$, $\vec{\Psi}_{G.b}, \vec{\vec{\Lambda}}_{G.b}$,and $\vec{\Sigma}_{G.b}$ accordingly as the companion form equivalents of $\alpha_{F.bt}, \Psi_{G.b}, \Lambda_{G.b}$,and $\Sigma_{G.b}$, respectively. We first run the Kalman filter forward to obtain the sequence $\left\{\vec{G}_{bt|t}\right\}$ that accounts for $\vec{\alpha}_{F.bt}$ and the corresponding covariance matrix $\vec{P}_{GbT|T}$ in period T based on all available sample information. This implies the following prediction and updating equations:

$$\begin{aligned}
\vec{G}_{bt+1|t} &= \vec{\alpha}_{F.bt} + \vec{\Psi}_{G.b}\vec{G}_{bt|t} \\
P_{Gbt+1|t} &= \vec{\Psi}_{G.b}P_{Gbt|t}\vec{\Psi}'_{G.b} + \vec{\Sigma}_{G.b} \\
\vec{G}_{bt|t} &= \vec{G}_{bt|t-1} + P_{Gbt|t-1}\vec{\tilde{\Lambda}}'_{G.b}\left(\vec{\tilde{\Lambda}}_{G.b}P_{Gbt|t-1}\vec{\tilde{\Lambda}}'_{G.b} + \Sigma_{X.b}\right)^{-1}\left(\tilde{X}_{bt} - \vec{\tilde{\Lambda}}_{G.b}\vec{G}_{bt|t-1}\right) \\
P_{Gbt|t} &= P_{Gbt|t-1} - P_{Gbt|t-1}\vec{\tilde{\Lambda}}'_{G.b}\left(\vec{\tilde{\Lambda}}_{G.b}P_{Gbt|t-1}\vec{\tilde{\Lambda}}'_{G.b} + \Sigma_{X.b}\right)^{-1}\vec{\tilde{\Lambda}}_{G.b}P_{Gbt|t-1}
\end{aligned}$$

We can then use the algorithm proposed in Carter and Kohn (1994) and Frühwirth-Schnatter (1994) to sample the block level factors G_b conditional on the sequence $\left\{\vec{G}_{bt|t}\right\}$, the data X_{bt} and the relevant parameters, again taking into account the information carried by the common factors F through the time-varying intercept $\vec{\alpha}_{F.bt}$. More details are given in Appendix A.

We assume the prior distribution for all factor loadings Λ and autocorrelation coefficients Ψ to be Gaussian with mean zero and variance one. The prior distribution for the variance parameters is that of an inverse chi-square distribution with ν degrees of freedom and a scale of d where ν and d^2 are set to 4 and 0.01, respectively. [2] After discarding the first 2,000 draws as a burn-in, we take another 25,000 draws, storing every 50-th draw. The reported statistics for posterior distributions are based on these 500 draws. Results obtained from storing every one of the first 8,000 draws after burn-in are very similar.

2.3 Related Work

Multilevel factor models have been considered extensively in the psychology literature. With the size of the panel being large in only one dimension and assuming a strict factor structure, these models can be estimated by maximum likelihood. See Goldstein and Browne (2002) for a review. However, these models do not allow for dynamics. Dynamic hierarchical linear models were considered by Gammerman and Mignon (1993), but there are no latent variables.

The two models closest to ours are Diebold et al. (2008) and Kose et al. (2003). Like us, Diebold et al. (2008) also assume a hierarchy of the factors. In their model, country-level yield curve factors share a common factor structure at the global level. In contrast to the one step estimation that we propose, Diebold et al. (2008) estimate their model in two steps. In the first step, they estimate the country level factors using classical OLS. In the second step, Diebold et al. (2008) treat the

[2] If θ is distributed as inverse χ^2 with ν degrees of freedom and a scale of d, written $\theta \sim I\chi^2(v, d^2)$, then θ is distributed as an inverse gamma with parameters $\alpha/2$ and $\beta/2$, where $\alpha = \nu$ and $\beta = d^2\nu$. We use this equivalence in our procedure and sample variance parameters based on the χ^2 distribution.

first step estimates as observable and use them as inputs to the measurement equation of a second state space model which they estimate using Gibbs sampling methods. This procedure has at least two disadvantages. First, it does not take into account the global factor dynamics in the estimation of the country-level block factors. Hence, the potentially important information embedded in the comovement across blocks (countries) is essentially ignored when estimating the block level factors. Second, the sampling uncertainty around the block level factors is not taken into account in the estimation of the global factors.

As shown above, the hierarchical structure that both our models have in common implies that the transition equation for the factors at a given level has a time varying intercept which depends on the factors at the next higher level. This means that after a simple modification of the transition equation, any method for estimating latent factors in linear state space models can be applied to hierarchical factor models of the form proposed in Diebold et al. (2008) and in our paper. As we show in Section 4 below, this is true regardless of the number of levels in the hierarchy. It is important to note that any filtering or sampling algorithm can be adapted to account for the dependency on the higher level factors. In Section 2.2 above, we discuss the slight modification that is necessary for the case of the forward filtering backward smoothing algorithm of Carter and Kohn (1994) and Frühwirth-Schnatter (1994) which is arguably one of the most prominent of such methods.

While the model in Diebold et al. (2008) and the one suggested here share the hierarchical factor structure, our approach is more general for the following reasons. First, since their model is tailored to yield curve estimation, Diebold et al. (2008) impose a tight parametric structure on the factor loadings at the block level. This is not the case in our model. Second, in contrast to our model Diebold et al. (2008) do not allow for lags in the observation equation and therefore cannot account for lead-lag relationships between the different block-level factors, a feature that we believe is quite important for many macroeconomic applications.

Our three level factor model also shares common features with an approach that has previously been suggested in the macroeconomic literature. Kose et al. (2003), see also Kose et al. (2008), use multi-level factor models to study international business cycle comovements. In contrast to our approach, they do not model a hierarchal structure of the factors. Instead, they simply assume that economic fluctuations in each country are attributed to three types of shocks: a world, a regional and a country-specific business cycle component. For each observable variable i in country b, they have

$$x_{bit} = c_i F_t + d_{bi} e_{Gbt} + e_{bit}$$

where F_t is a world factor, e_{Gbt} is a common shock specific to region b, and where e_{bit} is a component specific to variable i in country b.[3]

Our hierarchical multilevel model differs from theirs in a number of ways. First and foremost, we take a 'bottom up' approach while theirs is 'top down'. This means that in our model, the level two factors in the form of G_{bt} are well defined, and could be objects of independent interest. In contrast, a top down approach only yields a block-level component e_{Gbt} that is orthogonal to F_t. This means in the context of a multi-country model that we can identify a Euro factor that is uncorrelated with a global factor and an Asia-specific factor, but the model does not deliver an estimate of the state of the Euro or the Asian economy.

There are other differences between our model and that of Kose et al. (2003). While their F_t and G_{bt} are scalars, we allow for multiple common and multiple block-level factors. Moreover, our model allows for lagged factor observations to enter the measurement equation at the various levels. In contrast, their model only allows contemporaneous factor observations to have an effect on time-t measurements. More specifically, the model by Kose et al. (2003) is most comparable to a simplified version of our model that restricts the lagged loadings to zero, as presented in (6). Then their loading on the world factor c_i plays the role of our $\lambda^i_{G.b}\lambda^j_{F.b}$ and their loading d_{bi} on the regional factor is our $\lambda_{G.bi}$. Since we impose the structure that G_{bt} is linear in F_t, the responses of shocks to F_t for all variables in block b can only differ to the extent that their exposure to the block-level factors differs, whereas c_i is unconstrained in Kose et al. (2003). We believe that this restriction is a sensible one in many economic applications. For example, it appears quite plausible to assume that the effect of an economy-wide shock is similar for the capacity utilization in different industries and is also similar for the measures of employment across different age groups, but that the two sets of variables show quite different loadings across groups.

By imposing the hierarchical structure, we have a total of $K_G \times K_F$ and $N \times K_G$ parameters characterizing loadings on F_t and G_t, whereas Kose et al. (2003) have $N \times K_F$ and $N \times K_G$ parameters, respectively. As K_G is much smaller than N, our framework is considerably more parsimonious and thus its estimation likely to be computationally more efficient in applications where N is large. In addition, our model allows for dynamic loadings, but their model restricts F_t

[3] A similar framework was recently used by Stock and Watson (2008a) to analyze national and regional factors in housing construction.

and e_{Gbt} to have non-zero loadings only contemporaneously. In this regard, our model is much less restrictive.

Because the factors and the loadings in the common component are not separately identifiable, a different modeling strategy as discussed earlier is to assume a hierarchy for the factor loadings. Lopes et al. (2008) and Calder (2007) specify the hyperparameters to take advantage of spatial distance between units. With economic data, the distance between units is often not well defined. More importantly, for real time monitoring of economic activity and business cycle analysis, tracking the factors at the block level (say, the housing) is of independent interest.

Also somewhat related to our model is the work of Francke and de Vos (2000). They consider a hierarchical model in which the common and block components are random walks. They directly model the trend components without distinguishing the factor from its loadings. Milani and Belviso (2006) organize a large panel of macroeconomic time series for the US into blocks of data. They do not assume the existence of comovement beyond the block structure, but instead model the dynamic evolution of the different block factors jointly within a VAR. Clearly, this approach imposes a constraint on the number of block factors that one can allow for. Hallin and Liska (2008) also study dynamic factor models with a block structure using dynamic principal components. In their analysis, the factors can fall into as many as 2^K possible categories, where K denotes the number of blocks. This can be computationally challenging if K turns out to be large. Our work is distinct from theirs, as by exploiting the timing of data release or economic and geographical structure, we assume that the block structure of the data is known.

In terms of estimation, Otrok and Whiteman (1998) estimate latent dynamic factors by considering their conditional joint distribution. The main practical limitation is that they have to invert a variance-covariance matrix of rank T at each iteration of their Gibbs sampling algorithm. Hence, estimation becomes computationally demanding for problems when N and T are both large. Our experimental models here have up to $N > 400$ series and $T > 200$, and we anticipate using as many as 1,000 series in a fully fledged analysis. To alleviate the dimensionality problem, we put more structure on the block factors G_t. This enables us to exploit the prediction error decomposition of F_t and G_t which avoids inverting large matrices.

An alternative to Gibbs sampling is to estimate G_t by principal components, and then estimate F_t from the principal components estimates of G_t. This method was implemented in Beltratti and Morana (2008). However, sequential estimation by principal components would not take into account the dependence of G_t on F_t through α_{Ft}. These 'unrestricted' estimates of G_t should thus be less efficient than our one step estimates. Another advantage of our approach is that the posterior distributions allow us to assess sampling uncertainty about the estimated factors. While the large sample theory for principal components estimation of G_t and F_t is given in Bai and Ng (2006), the properties of the principal components estimator for F_t based upon a first step estimation of G_t by principal components is not known. It remains unclear how to obtain theoretical prediction intervals or assess the sampling uncertainty of counter-factual analysis within the two-step principal components framework.

Our model is unique in that it permits information aggregation while preserving the heterogeneous characteristics of the data at the block level in an internally coherent manner. Nonetheless, as a cross-check, it is useful to compare the estimates produced by our three level model with those obtained from principal components analysis. Hereafter, we use a 'tilde' to denote estimates obtained by the method of principal components, and a 'hat' to denote estimates obtained from our MCMC algorithm. That is, $\widehat{G}_t$ denote the posterior means of the block-level dynamic factors while $\widehat{F}_t$ are the posterior means of the factors common to G_t. In contrast, we refer to $\widetilde{F}_t$ as the principal component estimates obtained using all data at once and let $\widetilde{F}_t(\widetilde{G}_t)$ denote the two step principal components estimates (obtained from extracting principal components from the block-level principal components estimates). However, it should be kept in mind that the method of principal components estimates the static factors, whereas we estimate the dynamic factors, which should generally be smoother than the static factors.

We use the principal components estimated for each block, denoted $\widetilde{G}_{b,PC}$, as starting values for G_t. The principal components extracted from the data pooled across blocks are then used as starting values for F_t. Note that the principal components only identify the factor space using the normalization that $\widetilde{\Lambda}'_{G,PC}\widetilde{\Lambda}_{G,PC}/N = I_r$ and the matrix $\widetilde{G}'_{PC}\widetilde{G}_{PC}$ is diagonal. We use alternative identification assumptions. Therefore, our starting values may be far from the true values. As a cross-check on our choice of initialization, we also run the MCMC algorithm using randomly generated numbers for the factors as starting values and find that the sampler converges to the same posterior means.[4]

[4]In an unreported Monte Carlo exercise, we estimate a model with three blocks of data. We treat the posterior means of the parameter estimates as well as $\widehat{F}_t$ and $\widehat{G}_t$ as 'true' values, and resample ϵ_X to a set of simulated data. The simulated data are then used to estimate the parameters. The estimated factors are close to the true factor processes as implied by the simulated data, and the posterior mean of the parameters are also close to the 'true' values. This is not reported to conserve space.

3 A Six Block-Three Level Model of Real Activity

In the applications considered in this paper, we use a balanced panel of monthly data from 1992:01-2009:07. The data are transformed to be stationary using Stock and Watson (2008b) as a guide. A complete list of the series used along with their source and the transformation applied is provided in the appendix. After the data transformation, our sample effectively starts in 1992:4, giving $T = 207$ observations for all blocks. Summary statistics based on factors estimated by principal components are reported in Table 1.

We organize the plethora of economic data for real economic activity in the US according to the various statistical releases in which they are published. By the construction of the data, these releases broadly correspond to economic categories. For example, the Bureau of Labor Statistics (BLS) publishes data on the employment situation in the first week of each month. Their release consists of two separate surveys: the establishment survey with about 90 series summarizing the state of the labor market from the employers' perspective, and the household report which collects about 80 series on the labor market conditions faced by employees. Assuming that both surveys contain useful information about the state of the labor market, we treat them as separate blocks of data and will be interested in the dynamics they share with the other output related blocks. The Board of Governors releases about 40 series of industrial production and 25 series of capacity utilization data in the third week of each month. Finally, the Census Bureau publishes about 60 series on manufacturers' shipments, inventories and orders of durable goods in its advance report on durable goods in the last week of the month. We complement these releases on output related economic categories with survey information on manufacturing activity. We collapse three such surveys into one block of data: the Institute for Supply Management's (ISM) survey, the Federal Reserve Bank of Philadelphia's Business Outlook, and the Federal Reserve Bank of Chicago's Midwest Manufacturing survey.

We estimate a three level dynamic hierarchical factor model for six blocks of data related to output in the US economy that are released at different dates in each month: industrial production (IP), capacity utilization (CU), the establishment survey (ES), the household survey (HS), manufacturers' surveys (MS), and durable goods (DG). An important aspect of our analysis is that we use prior information to identify the factors. This involves grouping the series in the dataset into blocks of variables, and then ordering the variables in each block so that the series thought most likely to be representative of comovement in a given block are put in positions one through K_{Gb}. This is listed below.

Block		N	Variable Ordered First	Variable Ordered Second
1	CU	25	Machinery	Motor Vehicles and Parts
2	IP	38	Durable Consumer Goods	Nondurable Consumer Goods
3	ES	82	All Employees: Wholesale Trade	Avg Wkly Earnings: Construction
4	HS	92	Civilian Labor Force: Men: 25-54 Years	Unemp. Rate Full-Time Men Workers
5	MS	35	PMI Composite Index	Phila FRB General Activity Index
6	DG	60	Inventories: Machinery	Mfrs' Unfilled Orders: Machinery

3.1 Results

According to the IC_2 criterion of Bai and Ng (2002), four of the six blocks have either one or two factors. However, the criterion suggests that the HS and MS blocks may have as many as eight factors. Although our Bayesian estimation approach generally allows for different numbers of factors across blocks, we let all blocks be driven by two block-level components so as to enhance comparability of the results. We assume one common factor at the aggregate level.[5] Our model is described by the following set of parameters : $K_F = 1$, $K_{Gb} = 2$ for all b, $l_{Fk} = 2$, $l_{Gb} = 2$, $q_{Fk} = q_{Gb.j} = q_{Xb.i} = 1$ for all $b = 1, \dots B$, $k = 1, \dots K_F$, $j = 1, \dots K_{Gb}$, and $i = 1, \dots N_b$. We note that the estimated factors and idiosyncratic errors are generally mildly persistent, suggesting that the transformed data used in the analysis are stationary.

The top panel of Table 2 reports the posterior means and standard errors of the dynamic parameters. The common factor has an autoregressive coefficient of .88. The block-level factors have varying degrees of persistence, and many of the block-level factors are close to white noise. The block-level shocks tend to have larger variance than the shocks to the common factors.

In this model, there are $N \times 2$ loadings on G_t, and $K_G \times 1$ loadings on F_t, where $K_G = 12$ and N=332. Instead of reporting all the loadings, we summarize the properties of the model by evaluating the relative importance of the common, block-level, and idiosyncratic variation.[6] The bottom panel of Table 2 shows that there is substantial heterogeneity across blocks. Of the six blocks considered, the CU, the IP, and the ES blocks have the largest common component, explaining about 30% of the variation in the data of the block. The block-level shocks roughly explain another 10 to 15% of the variation in these three blocks. Thus, the common and block-level factors in our sample of economic variables explain close to 40% of the variation in the blocks. This is similar to what one finds in principal components analysis applied to the much analyzed Stock and Watson dataset with 132 series, where the first five factors are found to explain about 40% of

[5]Initial estimation assuming two common factors suggests that the second factor has a very small variance, and dropping it did not lead to any noticeable change in the decomposition of variance.

[6]To preserve space, we present the average shares across all series in a given block.

the data.

While aggregate shocks to the CU, IP, and ES blocks are more important than the block-level shocks, the block-level component is larger than the common component in all remaining blocks. Shocks common to the MS block account for around 22% of the variations, compared to the common component of about 12%. The result that stands out in Table 2 is that the idiosyncratic component always explains the largest share of variation. In particular, 77% of the variation in the Household Survey block is idiosyncratic, and only 8% of the variation in that block is explained by the common factor F. These results highlight the difficulty in distilling information relevant for aggregate policy from observed data, as block-level information can be disguised as common variations, and a large idiosyncratic component can make precise estimation of the common factor space difficult.

As noted earlier, if block-level variations are important, some of the principal components extracted from the entire panel of data might correspond to block-level factors. To investigate this issue, we regress the principal components $\widetilde{F}_{kt}$ on $\widehat{F}_t$ to obtain residuals $\widetilde{e}_{kt}$ for each $k = 1, \ldots K_F$. These are variations deemed common by the method of principal components but not by our $\widehat{F}_t$. We then check if these residuals can be explained by our estimated block-level factors by regressing $\widetilde{e}_{kt}$ on $\widehat{G}_{bjt}$. To conserve space, Table 3 reports only the values of R^2 that exceed 10%. Evidently, many of the block-level variations are correlated with the factors estimated by the method of principal components from the entire data panel. The first and second principal components are correlated with variations in the Establishment Survey block ($b = 3$) with a correlation as high as 0.32, while the fourth principal component is correlated with the Household Survey block ($b = 4$). This could be a consequence of the fact that the employment block constitutes almost one third of the data, and common variations in the Household Survey block are deemed more important in principal component analysis than in our framework. The factors of the Capacity Utilization ($b = 1$) and Industrial Production blocks ($b = 2$) are strongly correlated with the fifth principal component. Overall, we interpret these results as suggesting that variations identified as common by principal component analysis may in fact occur at the block-level and not be genuinely common.

Figure 1 graphs the factors estimated using the different approaches. $\widehat{F}_t$ denotes the common factor estimated using our hierarchical model while $\widetilde{F}_t$ is the first principal component extracted from the from the entire data panel. Note that our $\widehat{F}_t$ is noticeably smoother than $\widetilde{F}_t$. In particular, the latter features large spikes in 1996 that are not prevalent in our common factor estimate $\widehat{F}_t$. One potential explanation for this relates to the government shutdown of the budget in January 1996. Due to the large number of employment related series in the dataset, the first principal component extracted from the panel puts a lot of weight on this block-level event. In contrast, it is appropriately treated as variations associated to a block of data using our hierarchical model. Notice further that the estimates nicely track the two recessions in our sample period. According to our common factor estimate, real activity bottomed at the end of 2001, consistent with the official NBER business cycle chronology which reports November 2001 as the trough of the recession. All estimates also document a collapse of real activity in late 2008 and a subsequent sharp rebound in early 2009.

4 A Four Level Hierarchical Dynamic Factor Model

Some blocks of data are naturally organized by subblocks. For example, the BLS combines the establishment and the household surveys into its employment report which is published once a month. While the time series in both surveys can be expected to share common variation, some of the dynamics might be specific to either of the two and not represent genuine common labor market related information. Similarly, different surveys of manufacturing activity are likely to feature both common and survey-specific components which might be important to disentangle in practice. Our hierarchical dynamic factor model can easily be extended to allow for a subblock level as we will discuss next.

We continue to let X_{bit} denote variables associated with block-level factors G_{bt}. To distinguish data associated with blocks that have subblocks from those that do not, let Z_{bsit} be the observed data for block b where s is an index for the subblocks. Let H_{bst} be the K_{Hbs} factors of subblock s in block b. Then a four level model can be represented by

$$\begin{aligned}
Z_{bsit} &= \lambda^i_{H.bs}(L)H_{bst} + e_{Zbsit}, \qquad \psi_{Z.bsi}(L)e_{Zbsit} = \epsilon_{Zbsit} \\
H_{bst} &= \Lambda_{G.bs}(L)G_{bt} + e_{Hbst}, \qquad \Psi_{H.bs}(L)e_{Hbst} = \epsilon_{Hbst} \\
G_{bt} &= \Lambda_{F.b}(L)F_t + e_{Gbt}, \qquad \Psi_{G.b}(L)e_{Gbt} = \epsilon_{Gbt} \\
\psi_{F.k}(L)F_{kt} &= \epsilon_{Fkt}.
\end{aligned}$$

The dependence of H_t on G_t implies that

$$H_{bst} = \alpha_{G.bst} + \Psi_{H.bs1}H_{bst-1} + \ldots + \Psi_{H.bsq_{Hbs}}H_{bst-q_{Hbs}} + \epsilon_{Hbst}$$

where $\alpha_{G.bst} = \Psi_{H.bs}(L)\Lambda_{G.b}(L)G_{bt}$. As in the three level model, the dependence of G_{bt} on F_t in turn implies

$$G_{bt} = \alpha_{F.bt} + \Psi_{G.b1}G_{bt-1} + \ldots + \Psi_{G.bq_{Gb}}G_{bt-q_{Gb}} + \epsilon_{Gbt}.$$

Conditional on G_{bt}, F_t, and Θ, we can draw H_{bst} for each s and b, and conditional on F_t, we can draw G_{bt} for each b.

Blocks that have a subblock structure can be combined with blocks that do not. A model with more levels can always be decomposed into a sequence of two level models. Of course, we will need to have a reasonable number of series at the subblock level. But conceptually, a model with 'branches' in some but not all blocks is straightforward to set up in our framework. Notice also that in contrast to multi-level factor models like the one in Kose et al. (2003), the hierarchical structure of our model ensures that allowing for subblocks does not increase the number of parameters to estimate with the order of variables included in the model.

Our hierarchical model setup is a dynamic factor model where each level admits a state-space representation that has a measurement and a transition equation. Hence, the estimation algorithm for a three level model discussed in section 2.2 can easily be modified to apply to a four level model.

4.1 A Four Level Model of Real Activity

We apply our four level hierarchical factor model to the same data used to estimate the three level model which we expand to include data related to the US housing market and data on retail as well as auto sales. We organize these data into fourteen subblocks each of which corresponds to a separate release by a statistical agency or other data provider. These are industrial production (IP), capacity utilization (CU), durable goods (DG), the establishment survey (ES), the household survey (HS), retail sales (RS), wholesale trade (WT), and auto sales (AUTO), housing starts (H-STARTS), new home sales (H-NEWSALES), existing home sales (H-EXISTSALES), the ISM manufacturing survey (ISM), the Philadelphia Fed Manufacturing survey (PHILAFED), and the Chicago Fed Midwest manufacturing survey (CHICFED). Figure 2 gives the time line of data releases of these fourteen subblocks for June 2009. We organize the fourteen subblocks in five blocks that together comprise 447 series. The first is an output block with subblocks IP, CU, and DG representing the goods production. The second is a labor market block consisting of subblocks ES and HS. The third is a demand block with subblocks RS, WT, and AUTO. The fourth is a housing market block which comprises the subblocks H-STARTS, H-NEWSALES and H-EXISTSALES. Finally, the fifth block is a manufacturing survey block with subblocks ISM, PHILAFED and CHICFED. We estimate one common factor ($K_F = 1$), and one common factor per block ($K_{Gb} = 1$), and one or two factors per subblock depending on the number of series they contain ($K_{Hbs} = 1$ or $K_{Hbs} = 2$). We interpret the estimated common factor as a factor for real economic activity. In the table below, we list each subblock along with the number of variables it contains, how many subblock factors we extract, and the variable that is ordered first.

Block	subblock	N	K_{Gb}	K_{Hbs}	Variable Ordered First
Production	CU	25	1	2	Capacity Utilization: Machinery
	IP	38	1	2	IP: Durable Consumer Goods
	DG	60	1	2	Manufacturers' Inventories: Machinery
Employment	ES	82	1	2	All Employees: Wholesale Trade
	HS	92	1	2	Civilian Labor Force: Men: 25-54 Years
Demand	RS	30	1	2	Retail Sales: General Merchandise Stores
	WS	54	1	2	Merchant Wholesalers: Sales: Automotive
	AUTO	4	1	1	Domestic Car Retail Sales
Housing	H-STARTS	24	1	2	Housing Starts: 1-Unit: West
	H-NEWSALES	5	1	1	New 1-Family Houses Sold: West
	H-EXISTSALES	4	1	1	NAR Total Existing Home Sales, Northeast
Mfg Surveys	ISM	9	1	1	ISM Mfg: PMI Composite Index
	PHILAFED	21	1	1	Phila FRB Bus Outlook: General Activity
	CHICFED	5	1	1	Chicago FRB: Midwest Manufacturing Index

As before, we assume the prior distribution for all factor loadings λ and autocorrelation coefficients ψ to be Gaussian with mean zero and variance one. The prior distribution for the variance parameters is that of an inverse chi-square distribution with ν degrees of freedom and a scale of d where ν and d^2 are set to 4 and 0.01, respectively.

Table 4 only reports the autoregressive parameters for G_t and F_t. As in the three level models considered in the previous section, the common factor is again more persistent than the block-level factors. The ψ_G for the output factor is close to that found for CU and IP in the three level model studied above, while that for the employment block is higher than that previously found for ES or HS. The demand factor is the least persistent of all block factors. Table 4 also reports the decomposition of variance, which is now performed at the subblock level. As in the three level model, idiosyncratic shocks dominate common variations at all levels of aggregation for the majority of subblocks. The only exceptions are H-EXISTSALES and CHICFED for which the share of variance explained by the subblock factor shocks exceeds the share of variance explained by idiosyncratic shocks. The CU and IP blocks continue to have the largest common component. However, this only accounts for about 10 % of the variation in these blocks. More generally, shocks at the block and the subblock level are more important drivers of time series variation in our panel than aggregate shocks in all subblocks.

Perhaps of most interest is an analysis of the state of real economic activity estimated with the model. This is presented in Figure 3. The solid line is the $\widehat{F}_t$ based on our model and the dotted lines are the estimated block factors $\widehat{G}_{bt}$, all standardized to have unit variance. This plot documents that while all subcomponents of output that we consider share the pattern of a strong collapse at the end of 2008 and a subsequent sharp rebound, the block dynamics were quite different. It is the flexibility of our model with respect to the lag structure of the factor loadings at the different levels of the hierarchy that allows us to disentangle these dynamics in a joint framework.

5 Real Time Monitoring

An important task for applied economists is to monitor economic activity in real time. This requires an updating of estimates of the state of the economy as new data become available. The hierarchical structure of our model makes it straightforward to update estimates of the factors at the different levels of the hierarchy once a new data release becomes available for a given set of economic indicators.

In particular, along with a draw from the posterior distribution of the model parameters and latent factors, a new time series observation on the series in a subblock can be used to update the subblock-specific factor H_{bs}. For the subblocks for which a new observation has not been observed yet, we can employ the model to generate linear forecasts. Together, the updated and predicted subblock factors can be used to update the estimate of the block-specific factor G_b. In a similar vein, we can predict the next observation for all block factors for which no new subblock information has come in. Along with the updated block-specific factor, we can then update our estimate of the economy-wide factor F. Hence, each time new information on any of the subblocks of data becomes available, we can update our estimate of the state of the economy through the hierarchical structure of our model. The exact algorithm used to monitor the factors at the different levels of the hierarchy is as follows:

1. If data for the month on subblock $s \in [1, S]$ in block b has been released, use data Z_{bsit} as well as a draw j from the posterior of the model parameters $\vec{\tilde{\Lambda}}_{H.bs}$,$\vec{\Psi}_{H.bs}$, $\vec{\Sigma}_{H.bs}$, and $\vec{\Sigma}_{Z.bs}$ and the model factors $\{H_{bst}\}_{t=1}^{T}$ to update H_{bsT+1}.[7] This is done as follows. We use the Kalman

[7]The arrowed parameters are the vectorized versions of the model factors and parameters introduced above. They are defined in detail in Section B in the appendix. As an example, $\vec{H}_{bst}$ denotes the stacked vector of observations $\left\{H_{bst}, H_{bst-1}, \ldots, H_{bst-l_H^*}\right\}$. Notice also that we drop the superscripts j denoting the individual draws from the posterior distribution of the model parameters in the outline of the monitoring algorithm so as to enhance readability.

filter to obtain the mean and variance of the conditional distribution $H_{bsT+1|T+1}$:

$$\begin{aligned}\vec{H}_{bsT+1|T+1} &= \vec{H}_{bsT+1|T} + K_{H.bsT}\left(\widetilde{Z}_{bsT+1} - \tilde{\vec{\Lambda}}_{H.bs}\vec{H}_{bsT+1|T}\right), \\ P_{H.bsT+1|T+1} &= P_{H.bsT+1|T} - K_{H.bsT}\tilde{\vec{\Lambda}}_{H.bs}P_{H.bsT+1|T},\end{aligned}$$

where

$$\begin{aligned}\vec{H}_{bsT+1|T} &= \vec{\alpha}_{G.bsT} + \vec{\Psi}_{H.bs}\vec{H}_{bsT|T}, \\ P_{H.bsT+1|T} &= \vec{\Psi}_{H.bs}P_{H.bsT|T}\vec{\Psi}'_{H.bs} + \vec{\Sigma}_{H.bs}, \\ K_{H.bsT} &= P_{H.bsT+1|T}\tilde{\vec{\Lambda}}'_{H.bs}\left(\tilde{\vec{\Lambda}}_{H.bs}P_{H.bsT+1|T}\tilde{\vec{\Lambda}}'_{H.bs} + \vec{\Sigma}_{Z.bs}\right)^{-1}.\end{aligned}$$

Given $\vec{H}_{bsT+1|T+1}$ and $P_{H.bsT+1|T+1}$, we generate a draw $\vec{H}_{bsT+1}$ from $N(\vec{H}_{bsT+1|T+1}, P_{H.bsT+1|T+1})$.

2. If data on subblock s has not been released for the month, we "predict" $\vec{H}_{bsT+1}$ and $P_{H.bsT+1|T}$ using

$$\begin{aligned}\vec{H}_{bsT+1|T} &= \vec{\alpha}_{G.bsT} + \vec{\Psi}_{H.bs}\vec{H}_{bsT|T}, \\ P_{H.bsT+1|T} &= \vec{\Psi}_{H.bs}P_{H.bsT|T}\vec{\Psi}'_{H.bs} + \vec{\Sigma}_{H.bs},\end{aligned}$$

and generate a draw $\vec{H}_{bsT+1}$ from $N(\vec{H}_{bsT+1|T}, P_{H.bsT+1|T})$.

3. Conditional on $\vec{H}_{bT+1} = \left(\vec{H}_{b1T+1}, \ldots, \vec{H}_{bST+1}\right)$, $\tilde{\vec{\Lambda}}_{G.b}$, $\vec{\Psi}_{G.b}$, $\vec{\Sigma}_{G.b}$ and $\vec{\Sigma}_{H.b}$, we update G_b. That is, we first use the Kalman filter to obtain the mean and variance of the conditional distribution $\vec{G}_{bT+1|T+1}$as above. Given $\vec{G}_{bT+1|T+1}$ and $P_{G.bT+1|T+1}$, we then generate a draw $\vec{G}_{bT+1}$ from $N(\vec{G}_{bT+1|T+1}, P_{G.bT+1|T+1})$.

4. If no data for block b has been released for the month yet, we "predict" $\vec{G}_{bT+1}$ and $P_{G.bT+1|T}$ using

$$\begin{aligned}\vec{G}_{bT+1|T} &= \vec{\alpha}_{F.bT} + \vec{\Psi}_{G.b}\vec{G}_{bT|T}, \\ P_{G.bT+1|T} &= \vec{\Psi}_{G.b}P_{G.bT|T}\vec{\Psi}'_{G.b} + \vec{\Sigma}_{G.b},\end{aligned}$$

and generate a draw $\vec{G}_{bT+1}$ from $N(\vec{G}_{bT+1|T}, P_{G.bT+1|T})$.

5. Conditional on $\vec{G}_{T+1} = \left(\vec{G}_{1T+1}, \ldots, \vec{G}_{BT+1}\right)$, $\tilde{\vec{\Lambda}}_F$, $\vec{\Psi}_F$, $\vec{\Sigma}_F$, and $\vec{\Sigma}_G$, we update F. Again, we first use the Kalman filter to obtain the mean and variance of the conditional distribution $F_{T+1|T+1}$.Then, given $\vec{F}_{T+1|T+1}$ and $P_{F.T+1|T+1}$, generate a draw $\vec{F}_{T+1}$ from $N(\vec{F}_{T+1|T+1}, P_{F.T+1|T+1})$.

Repeating this procedure for many draws from the posterior distribution of the parameters and model factors provides us with a distribution of factor estimates each time a new data point is observed. Thus, if five blocks of data with a total of fourteen subblocks are released successively over the course of the month, we get fourteen updates of our common factor F_{T+1} that captures the pervasive comovement across all blocks of data.

In related work, Giannone, Reichlin, and Small (2008) use a factor model to forecast quarterly GDP growth as monthly information arrives in real time. They estimate their model using a two-step approach that combines principal components with Kalman filtering techniques. Banbura et al. (2008) extend the model in Giannone, Reichlin, and Small (2008) to the components of GDP by incorporating temporal aggregation and accounting identities. While these two papers are similar to ours in the sense that they also employ a large cross-section of macroeconomic time series in a real-time monitoring exercise, they do not take into account the block structure of the data explicitly which is a unique feature of our modeling approach.

5.1 Monitoring the 2007-2009 Economic Downturn

We illustrate how our model can be used to track the state of the economy while new data becomes available based on the four level model of real activity presented in Section 4.1. In particular, we document the evolution of the factor estimates over the course of the last two years beginning in July 2007. The starting values for the monitoring exercise are the estimates obtained using data from April 1992 through July 2007. We contrast our continuous update which is obtained each time a new observation on any subblock of the dataset becomes available with a monthly update of the model which we obtain at the end of each month when a full set of new observations on all 447 series in our panel is available.[8]

Figure 4 shows the estimated posterior means of the continuous update and the monthly update for the common factor F. The latter has the form of a step function as it only features a new data point at the end of each month. According to this plot, the continuous updates of activity appear to provide a more timely signal about the state of the economy than the monthly updates. Especially around the collapse of real activity in the fall of 2008 and its subsequent rebound in 2009, the continuous update is consistently leading the monthly update. This highlights the usefulness of our hierarchical model for tracking the state of the economy when data on different economic categories

[8] In this exercise, we use final data releases instead of actual real time data which are not available for the large cross-section of macroeconomic time series that we consider. We are currently building such a database. To the extent that data revision errors are series specific and don't share a common component, the factors extracted from revised and real-time data should be very similar.

become available in a consecutive way. This is particularly important for real time policy analysis which heavily relies on the timeliness of signals about changes in the course of economic growth.

Figures 5 to 7 show the continuous and monthly updates for the estimated block factors $\widehat{G}$ of the Demand, Housing, and Manufacturing Survey blocks, respectively. All show a very similar pattern: the intra-month update is consistently leading the monthly update throughout the last recession. Hence, in addition to tracking aggregate economic activity in real time, our model also allows us to update estimates of factors common to blocks of data while new data on subblocks become available. Since in our application, these blocks of data correspond to economic categories, the estimated block factors are objects of independent interest. In this respect, it is important to note that the exact timing of the downturn and following rebound differs across the various blocks of data. For example, while the demand and the manufacturing survey blocks appears to have bottomed in December 2008, housing activity growth as captured by the housing block factor only picked up in March 2009.[9] These results highlight that business cycle variations are only loosely synchronized across the different sectors of the economy. It is therefore important to allow for lead-lag relationships between the common factors and the factors capturing block-specific information. This is a distinctive feature that our model offers.

6 Conclusion

This paper lays out a framework for analyzing dynamic hierarchical factor models. The approach has three advantages. First, by extracting common components from blocks, the estimated factors have a straightforward interpretation. Explicitly modeling the block-level variation also resolves an important drawback of standard (two level) factor models in which common shocks at the block-level can be confounded with genuinely common shocks. Second, the blocks can be defined to take advantage of the timing of data releases, which makes the framework suitable for real time monitoring of economic activity. Third, the framework allows for a more disaggregated analysis of economic fluctuations while still achieving a reasonable level of dimension reduction. While a two level model only enables counter-factual analyses of aggregate or idiosyncratic shocks, the effects of aggregate, block-level, and idiosyncratic shocks can be coherently analyzed in our framework.

[9]Note that since the data have been transformed into growth rates for most blocks, the factor estimates capture dynamics in the growth cycle rather than the level cycle.

References

Aguilar, G. and West, M. 2000, Bayesian Dynamic Factor Models and Portfolio Allocation, *Journal of Business and Economic Statistics* **18**, 338–357.

Bai, J. and Ng, S. 2002, Determining the Number of Factors in Approximate Factor Models, *Econometrica* **70:1**, 191–221.

Bai, J. and Ng, S. 2006, Confidence Intervals for Diffusion Index Forecasts and Inference with Factor-Augmented Regressions, *Econometrica* **74:4**, 1133–1150.

Banbura, M., Angelini, E. and Runstler, G. 2008, Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model, ECB WP 953.

Beltratti, A. and Morana, C. 2008, International Shocks and National House Prices, Bocconi University.

Brooks, R. and DelNegro, M. 2006, Firm-Level Evidence on International Stock Market Comovement , *Review of Finance* **10:1**, 69–98.

Calder, C. 2007, Dynamic Factor Process Convolution Models for Multivariate Space-Time Data with Application to Air Quality Assessment, *Environmental and Ecological Statistics* **14**, 229–247.

Carter, C. K. and Kohn, R. 1994, On Gibbs Sampling for State Space Models, *Biometrika* **81:3**, 541–533.

Diebold, F., Li, C. and Yue, V. 2008, Global Yield Curve Dynamics and Interations: A Dynamic Nelson-Siegel Approach, *Journal of Econometrics* **146**, 315–363.

Francke, M. and de Vos, A. 2000, Efficeint Computation of Hierarchical Trends, *Journal of Business and Economic Statistics* **18:1**, 51–57.

Frühwirth-Schnatter, S. 1994, Data Augmentation and Dynamic Linear Models, *Journal of Time Series Analysis* **15**, 183–202.

Gammerman, D. and Mignon, H. S. 1993, Dynamic Hierarchical Models, *Journal of Royal Statistical Society Series B* **55**, 629–642.

Geweke, J. and Zhou, G. 1996, Measuring the Pricing Error of the Arbitrage Pricing Theory, *Review of Financial Studies* **9:2**, 557–87.

Giannone, D., Reichlin, L. and Small, D. 2008, Nowcasting: The Real-Time Informational Content of Macroeconomic Data, *Journal of Monetary Economics* **55**, 665–676.

Glick, R. and Rogoff, K. 1995, Global versus Country-Specific Productivity Shocks and the Current Account, *Journal of Monetary Economics* **35**, 159–192.

Goldstein, H. and Browne, W. 2002, Miltilevel Factor Analysis Using Markov Chain Monte Carlo Estimation, *Latent variable and Latent Structure Models*, pp. 225–243.

Hallin, M. and Liska, R. 2008, Dynamic Factors in the Presence of Block Structure, European University Institute WP 2008/22.

Kim, C. and Nelson, C. 2000, *State Space Models with Regime Switching*, MIT Press.

Kose, A., Otrok, C. and Whiteman, C. 2003, International Business Cycles: World Region and Country Specific Factors, *American Economic Review* **93:4**, 1216–1239.

Kose, A., Otrok, C. and Whiteman, C. 2008, Understanding the Evolution of World Business CYcles, *International Economic Review* **75**, 110–130.

Lopes, H. and West, M. 2004, Bayesian Model Assessment in Factor Analysis, *Statistical Sinica* **14**, 41–87.

Lopes, H., Salazar, E. and Gamerman, D. 2008, Spatial Dynamic Factor Model, *Bayesian Analysis* **3:4**, 759–792.

Milani, F. and Belviso, F. 2006, Structural Factor-Augmented VARs and the Effects of Moentary Policy, *Topics in Macroeconomics* **6:3**, Article 2.

Mouchart, M. and Martin, E. S. 2003, Specification and Identification Issues in Models Involving a Latent Hierarhical Structure, *Journal of Statistical Planning and Inference* **111**, 143–163.

Otrok, C. and Whiteman, C. 1998, Bayesian Leading Indicators: Measuring and Predicting Economic Conditions in Iowa, *International Economic Review* **39:4**, 997–1014.

Stock, J. H. and Watson, M. W. 2006, Forecasting with Many Predictors, *Handbook of Forecasting*, North Holland.

Stock, J. H. and Watson, M. W. 2008a, The Evolution of National and Regional Factors in U.S. Housing Construction, Princeton University.

Stock, J. H. and Watson, M. W. 2008b, Forecasting in Dynamic Factor Models Subjet to Structural Instability, Princeton University.

Appendix

A The Three Level Model in Matrix Form

Let $l_{Gb} = (l_{Gb.1}, \dots l_{Gb.N_b})$ is a vector. Similarly, $q_{Xb} = (q_{Xb.1}, \dots q_{Xb.N_b})$ and $q_{Gb} = (q_{Gb.1}, \dots q_{Gb.K_{Gb}})$ are also vectors with possibly non-identical entries. Let $l_F = \max_{b \in B} l_{Fb}$, $l_G = \max_b(\max_{i \in N_b} l_{Gb.i})$, $q_X = \max_b(\max_{i \in N_b} q_{Xbi})$, and $q_G = \max_b(\max_{k \in K_{Gb}} q_{Gb.k})$.

Stacking up the data by blocks and letting

$$\begin{aligned} X_t &= \begin{pmatrix} X_{1t} & X_{2t} & \dots X_{Bt} \end{pmatrix}' \\ G_t &= \begin{pmatrix} G_{1t} & G_{2t} & \dots G_{Bt} \end{pmatrix}', \end{aligned}$$

we have

$$\begin{aligned} X_t &= \Lambda_G(L) G_t + e_{Xt} \\ G_t &= \Lambda_F(L) F_t + e_{Gt} \\ \Psi_F(L) F_t &= \epsilon_{Ft} \\ \Psi_X(L) e_{Xt} &= \epsilon_{Xt} \\ \Psi_G(L) e_{Gt} &= \epsilon_{Gt}. \end{aligned}$$

Then $\Lambda_G(L)$ is a $N \times K_G$ matrix polynomial of order l_G, $\Lambda_F(L)$ is a $K_G \times K_F$ matrix polynomial of order l_F, $\Psi_X(L)$ is a $N \times N$ matrix polynomial of order q_X, $\Psi_G(L)$ is a $K_G \times K_G$ matrix polynomial of order q_G, $\Psi_F(L)$ is a $K_F \times K_F$ matrix polynomial of order q_F. Finally, $\Sigma_X = diag(\sigma^2_{X11}, ..., \sigma^2_{XBN_B})$, $\Sigma_G = diag(\sigma^2_{G11}, ..., \sigma^2_{GBk_B})$, and $\Sigma_F = diag(\sigma^2_{F1}, ..., \sigma^2_{FK_F})$ are matrices of dimension $N \times N, K_G \times K_G$, and $K_F \times K_F$, respectively. To ensure identification of the block-level factors G, we assume that for $l = 0, \dots, l_G$

$$\Lambda_{G.l} = \begin{bmatrix} \Lambda_{G.1l} & 0 & \cdots & 0 \\ 0 & \Lambda_{G.2l} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Lambda_{G.Bl} \end{bmatrix}.$$

Sampling $\{F_t\}$

In companion form as

$$\begin{pmatrix} F_t \\ F_{t-1} \\ \vdots \\ F_{t-q_F+1} \end{pmatrix} = \begin{bmatrix} \Psi_{F.1} & \Psi_{F.2} & \cdots & \Psi_{F.q_F} \\ I & 0 & & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} \begin{pmatrix} F_{t-1} \\ F_{t-2} \\ \vdots \\ F_{t-q_F} \end{pmatrix} + \begin{pmatrix} \epsilon_{Ft} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

or

$$\vec{F}_t = \vec{\Psi}_F \vec{F}_{t-1} + \vec{\epsilon}_{Ft}$$

To obtain estimates of the global factors F given the block factors G, we have to perform the following steps. First, pre-whiten the observation equation

$$G_t = \Lambda_F(L) F_t + e_{Gt}$$

so that its errors are i.i.d. This gives $\Psi_G(L)G_t = \Psi_G(L)\Lambda_F(L)F_t + \epsilon_{Gt}$ or

$$\widetilde{G}_t = \widetilde{\Lambda}_F(L)F_t + \epsilon_{Gt}$$

where $\widetilde{G}_t = \Psi_G(L)G_t,$ and where $\widetilde{\Lambda}_F(L) = \Psi_G(L)\Lambda_F(L) = \widetilde{\Lambda}_{F0}+\widetilde{\Lambda}_{F1}L+...+\widetilde{\Lambda}_{Fs^*_F}L^{l^*_F}$ is a $K_G\times K_F$ matrix polynomial of order $l^*_F = q_G + l_F$. Stacking the lags of F, this gives the companion form:

$$\widetilde{G}_t = \begin{bmatrix} \widetilde{\Lambda}_{F.0} & \widetilde{\Lambda}_{F.1} & \cdots & \widetilde{\Lambda}_{F.l^*_F} \end{bmatrix} \begin{pmatrix} F_t \\ F_{t-1} \\ \vdots \\ F_{t-l^*_F} \end{pmatrix} + \epsilon_{Gt}$$

$$\begin{pmatrix} F_t \\ F_{t-1} \\ \vdots \\ F_{t-l^*_{F+1}} \end{pmatrix} = \begin{bmatrix} \Psi_{F.1} & \cdots & \Psi_{F.q_F} & 0 & \cdots & 0 \\ I & 0 & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \vdots & & \vdots \\ 0 & \cdots & I & 0 & \cdots & 0 \end{bmatrix} \begin{pmatrix} F_{t-1} \\ F_{t-2} \\ \vdots \\ F_{t-l^*_F} \end{pmatrix} + \begin{pmatrix} \epsilon_{Ft} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

or

$$\widetilde{G}_t = \vec{\widetilde{\Lambda}}_F\vec{F}_t + \epsilon_{Gt} \qquad \text{and } \vec{F}_t = \vec{\Psi}_F\vec{F}_{t-1} + \vec{\epsilon}_{Ft}$$

where $\vec{\Sigma}_F = Var(\vec{\epsilon}_{Ft}) = \begin{pmatrix} \Sigma_F & 0 \\ 0 & 0 \end{pmatrix}$.

Denote Ξ_F the set of parameters $\{\vec{\widetilde{\Lambda}}_F, \vec{\Psi}_F,\Sigma_G,\vec{\Sigma}_F\}$. Then, following Carter and Kohn (1994), the conditional distribution of the factors $\vec{F}$ given the pre-whitened block factors $\{\widetilde{G}_t\}$ and the parameters Ξ_F can be obtained by performing the following steps. First run the Kalman filter forward to obtain estimates $\vec{F}_{T|T}$ of the (stacked) factors and their variance covariance matrix $\vec{P}_{T|T}$ in period T based on all available sample information:

$$\begin{aligned} \vec{F}_{t+1|t} &= \vec{\Psi}_F\ \vec{F}_{t|t} \\ \vec{P}_{Ft+1|t} &= \vec{\Psi}_F\vec{P}_{Ft|t}\vec{\Psi}'_F + \vec{\Sigma}_F \\ \vec{F}_{t|t} &= \vec{F}_{t|t-1} + \vec{P}_{Ft|t-1}\vec{\widetilde{\Lambda}}'_F\left(\vec{\widetilde{\Lambda}}_F\vec{P}_{Ft|t-1}\vec{\widetilde{\Lambda}}'_F + \Sigma_G\right)^{-1}\left(\widetilde{G}_t - \vec{\widetilde{\Lambda}}_F\vec{F}_{t|t-1}\right) \\ \vec{P}_{Ft|t} &= \vec{P}_{Ft|t-1} - \vec{P}_{Ft|t-1}\vec{\widetilde{\Lambda}}'_F\left(\vec{\widetilde{\Lambda}}_F\vec{P}_{Ft|t-1}\vec{\widetilde{\Lambda}}'_F + \Sigma_G\right)^{-1}\vec{\widetilde{\Lambda}}_F\vec{P}_{Ft|t-1} \end{aligned}$$

Next, draw $\vec{F}_T$ from its conditional distribution given Ξ_F and the data through period T:

$$\vec{F}_T|\{\widetilde{G}_t\}, \Xi_F \sim N(\vec{F}_{T|T}, \vec{P}_{FT|T})$$

Then, for $t{=}T{-}1,\ldots,1$ proceed backwards to generate draws $\vec{F}_{t|T}$ from

$$\begin{aligned} \vec{F}_{t|T}|\vec{F}^*_{t+1}, \{\widetilde{G}_T\}, \Xi_F &\sim N(\vec{F}_{t|t,\vec{F}^*_{t+1}}, \vec{P}_{t|t,\vec{F}^*_{t+1}}) \\ \text{where } \vec{F}_{t|t,\,\vec{F}^*_{t+1}} &= \vec{F}_{t|t} + \vec{P}_{t|t}\vec{\Psi}^{*\prime}_F\,(\vec{\Psi}^*_F\ \vec{P}_{t|t}\ \vec{\Psi}^{*\prime}_F + \Sigma_F)^{-1}\,(\vec{F}^*_{t+1} - \vec{\Psi}^*_F\ \vec{F}_{t|t}) \\ \text{and } \vec{P}_{t|t,\vec{F}^*_{t+1}} &= \vec{P}_{t|t} - \vec{P}_{t|t}\ \vec{\Psi}^{*\prime}_F\,(\vec{\Psi}^*_F\ \vec{P}_{t|t}\ \vec{\Psi}^{*\prime}_F + \Sigma_F)^{-1}\ \vec{\Psi}^*_F\ \vec{P}_{t|t}. \end{aligned} \qquad (11)$$

where $\vec{F}^*_t$ and $\vec{\Psi}^*_F$ are the first K_F rows of $\vec{F}_t$ and $\vec{\Psi}_F$, respectively.

Sampling $\{G_t\}$

For the block-level factors, (7) and (8) imply that

$$\Psi_G(L)G_t = \Psi_G(L)\Lambda_F(L)F_t + \epsilon_{Gt}.$$

This leads to the block-level transition equation

$$\begin{pmatrix} G_t \\ G_{t-1} \\ \vdots \\ G_{t-q_G+1} \end{pmatrix} = \begin{pmatrix} \alpha_{Ft} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{bmatrix} \Psi_{G.1} & \Psi_{G.2} & \cdots & \Psi_{G.q_G} \\ I & 0 & & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} \begin{pmatrix} G_{t-1} \\ G_{t-2} \\ \vdots \\ G_{t-q_G} \end{pmatrix} + \begin{pmatrix} \epsilon_{Gt} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

or

$$\vec{G}_t = \vec{\alpha}_{Ft} + \vec{\Psi}_G \vec{G}_{t-1} + \vec{\epsilon}_{Gt}$$

where

$$\alpha_{Ft} = \Psi_G(L)\Lambda_F(L)F_t.$$

A similar algorithm can be used to sample the block factors G. Since the block-dynamics are assumed to be independent, this can be done block by block. Recall that $\widetilde{X}_{bt} = \widetilde{\Lambda}_{G.b}(L)G_{bt} + \epsilon_{Xbt}, \forall\ b = 1, \ldots, B$, where $\widetilde{X}_{bt} = \Psi_{X.b}(L)X_{bt}$ and $\widetilde{\Lambda}_{G.b}(L) = \Psi_{X.b}(L)\Lambda_{G.b}(L)$ is a $N_b \times K_{Gb}$ matrix polynomial of order $l_G^* = q_X + l_G$. Furthermore, $G_{bt} = \alpha_{F.bt} + \Psi_{G.b1}G_{bt-1} + \ldots + \Psi_{G.bq_{Gb}}G_{bt-q_{Gb}} + \epsilon_{Gbt}$ where $\alpha_{F.bt} = \Psi_{G.b}(L)\Lambda_F(L)F_t, \forall\ b = 1, \ldots, B$. Together, these two equations imply the following state-space form

$$\widetilde{X}_{bt} = \begin{bmatrix} \widetilde{\Lambda}_{G.b0} & \widetilde{\Lambda}_{G.b1} & \cdots & \widetilde{\Lambda}_{G.bl_G^*} \end{bmatrix} \begin{pmatrix} G_{bt} \\ G_{bt-1} \\ \vdots \\ G_{bt-l_G^*} \end{pmatrix} + \epsilon_{Xbt}$$

$$\begin{pmatrix} G_{bt} \\ G_{bt-1} \\ \vdots \\ G_{bt-l_G^*} \end{pmatrix} = \begin{pmatrix} \alpha_{F.bt} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{bmatrix} \Psi_{G.b1} & \cdots & \Psi_{G.bq_G} & 0 & \cdots & 0 \\ I & 0 & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \vdots & & \vdots \\ 0 & \cdots & I & 0 & \cdots & 0 \end{bmatrix} \begin{pmatrix} G_{bt-1} \\ G_{bt-2} \\ \vdots \\ G_{bt-l_G^*-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{Gbt} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

or

$$\widetilde{X}_{bt} = \vec{\widetilde{\Lambda}}_{G.b}\vec{G}_{bt} + \epsilon_{Xbt} \qquad \text{and} \qquad \vec{G}_{bt} = \vec{\alpha}_{F.bt} + \vec{\Psi}_{G.b}\vec{G}_{bt-1} + \vec{\epsilon}_{Gbt}$$

where $\vec{\Sigma}_{G.b} = Var(\vec{\epsilon}_{Gbt}) = \begin{pmatrix} \Sigma_{G.b} & 0 \\ 0 & 0 \end{pmatrix}$.

Denote Ξ_{Gb} the set of parameters $\{\vec{\Lambda}_{G.b}, \vec{\Psi}_{G.b}, \vec{\Sigma}_{G.b}, \Sigma_{X.b}\}$. Conditional on Ξ_{Gb} and $\{F_t\}$, the above equations represent a state-space system with a time-varying intercept. We therefore need to slightly adjust the Carter and Kohn (1994) method as laid out before. The complete set of equations is as follows.

First, run the Kalman filter forward to obtain estimates $\vec{G}_{bT|T}$ of the factors and their variance covariance matrix $\vec{P}_{bT|T}$ in period T based on all available sample information. With the time-varying intercept $\vec{\alpha}_{F.bt}$, this implies the following steps:

$$
\begin{aligned}
\vec{G}_{bt+1|t} &= \vec{\alpha}_{F.bt} + \vec{\Psi}_{G.b}\vec{G}_{bt|t} \\
\vec{P}_{Gbt+1|t} &= \vec{\Psi}_{G.b}\vec{P}_{Gbt|t}\vec{\Psi}'_{G.b} + \vec{\Sigma}_{G.b} \\
\vec{G}_{bt|t} &= \vec{G}_{bt|t-1} + \vec{P}_{Gbt|t-1}\vec{\widetilde{\Lambda}}'_{G.b}\left(\vec{\widetilde{\Lambda}}_{G.b}\vec{P}_{Gbt|t-1}\vec{\widetilde{\Lambda}}'_{G.b} + \Sigma_{X.b}\right)^{-1}\left(\widetilde{X}_{bt} - \vec{\widetilde{\Lambda}}_{G.b}\vec{G}_{bt|t-1}\right) \\
\vec{P}_{Gbt|t} &= \vec{P}_{Gbt|t-1} - \vec{P}_{Gbt|t-1}\vec{\widetilde{\Lambda}}'_{G.b}\left(\vec{\widetilde{\Lambda}}_{G.b}\vec{P}_{Gbt|t-1}\vec{\widetilde{\Lambda}}'_{G.b} + \Sigma_{X.b}\right)^{-1}\vec{\widetilde{\Lambda}}_{G.b}\vec{P}_{Gbt|t-1}
\end{aligned}
$$

The Kalman filter iterations provide us with the conditional distribution of $\vec{G}_{bT|T}$ given Ξ_{Gb} and the data through period T:

$$
\vec{G}_{bT}|\{\widetilde{X}_{bt}\}, \Xi_{Gb} \sim N(\vec{G}_{bT|T}, \vec{P}_{GbT|T})
$$

Using again the algorithm of Carter and Kohn, we sample the entire set of factor observations conditional on the parameters Ξ_{Gb} and all the data. Given the Gaussianity and Markovian structure of the state-space model, the distribution of $\vec{G}_{bt}$ given $\vec{G}_{bt+1}$ and $\widetilde{X}_{bt}$is normal:

$$
\vec{G}_{bt}|\widetilde{X}_{bt}, \vec{G}^*_{bt+1}, \Xi_{Gb} \sim N(\vec{G}_{bt|t,\vec{G}^*_{bt+1}}, \vec{P}_{Gbt|t,\vec{G}^*_{bt+1}}) \tag{12}
$$

where

$$
\begin{aligned}
\vec{G}_{bt|t,\vec{G}^*_{bt+1}} &= E[\vec{G}_{bt}|\widetilde{X}_{bt}, \vec{G}^*_{bt+1}] \\
&= \vec{G}_{bt|t} + \vec{P}_{Gbt|t}\vec{\Psi}^{*\prime}_{G.b}\left(\vec{\Psi}^*_{G.b}\vec{P}_{Gbt|t}\vec{\Psi}^{*\prime}_{G.b} + \Sigma_{G.b}\right)^{-1}(\vec{G}^*_{bt+1} - \vec{\alpha}_{F.bt+1} - \vec{\Psi}^*_{G.b}\vec{G}_{bt|t}) \\
\vec{P}_{Gbt|t,\vec{G}^*_{bt+1}} &= Var(\vec{G}_{bt}|\widetilde{X}_{bt}, \vec{G}^*_{bt+1}) \\
&= \vec{P}_{Gbt|t} - \vec{P}_{Gbt|t}\vec{\Psi}^{*\prime}_{G.b}\left(\vec{\Psi}^*_{G.b}\vec{P}_{Gbt|t}\vec{\Psi}^{*\prime}_{G.b} + \Sigma_{G.b}\right)^{-1}\vec{\Psi}^*_{G.b}\vec{P}_{Gbt|t}
\end{aligned}
$$

where $\vec{G}^*_{bt+1}$ and $\vec{\Psi}^*_{G.b}$ denote the first K_{Gb} rows of $\vec{G}_{bt+1}$ and $\vec{\Psi}_{G.b}$, respectively. Given these conditional distributions, we can then proceed backwards to generate draws $\vec{G}^*_{bt}$ for $t = T-1, \ldots, 1$.

Decomposition of Variance Given the state space representation of the model, it is not hard to see that for each individual variable X_{bi},

$$vec(Var(X_{bi})) = \gamma'_{F.bi} vec(Var(F)) + \gamma'_{G.bi} vec(Var(e_{Gb})) + vec(Var(e_{Xbi})) \quad (13)$$

where

$$\begin{aligned}
\gamma'_{F.bi} &= \left(\sum_{l=0}^{l_G} \lambda'_{G.bil} \otimes \lambda'_{G.bil} \right) \cdot \left(\sum_{l=0}^{l_F} \Lambda_{F.bl} \otimes \Lambda_{F.bl} \right) \\
\gamma'_{G.bi} &= \left(\sum_{l=0}^{l_G} \lambda'_{G.bil} \otimes \lambda'_{G.bil} \right) \\
vec(Var(F)) &= \left[I - \sum_{q=1}^{q_F} (\Psi_{F.q} \otimes \Psi_{F.q}) \right]^{-1} \otimes vec(\Sigma_F) \\
vec(Var(e_{Gb})) &= \left[I - \sum_{q=1}^{q_G} (\Psi_{G.bq} \otimes \Psi_{G.bq}) \right]^{-1} \cdot vec(\Sigma_{G.b}) \\
vec(Var(e_{Xbi})) &= \left[1 - \sum_{q=1}^{q_X} \psi^2_{X.biq} \right]^{-1} \cdot \sigma^2_{X.bi}.
\end{aligned}$$

B The Four Level Model

Stacking all variables Z_{bsit} in a subblock and pseudo-differencing the serially correlated idiosyncratic components e_{Zbsit},the observation equation at the subblock level can be written as

$$\widetilde{Z}_{bst} = \widetilde{\Lambda}_{H.bs}(L)H_{bst} + \epsilon_{Zbst}, \qquad \forall\ b = 1, \ldots, B, \forall\ s = 1, \ldots, S,$$

where $\widetilde{Z}_{bst} = \Psi_{Z.bs}(L)Z_{bst}$ and where $\widetilde{\Lambda}_{H.bs}(L) = \Psi_{Z.bs}(L)\Lambda_{H.bs}(L)$ is a $N_b \times K_{Hbs}$ matrix polynomial of order $l_H^* = q_Z + l_H$. Moreover, the state equation at the subblock level is

$$H_{bst} = \alpha_{G.bst} + \Psi_{H.bs1}H_{bst-1} + ... + \Psi_{H.bsq_H}H_{bst-q_H} + \epsilon_{Hbst}$$

where

$$\alpha_{G.bst} = \Psi_{H.bs}(L)\Lambda_{G.b}(L)G_{bt}, \qquad \forall\ b = 1, \ldots, B, \forall\ s = 1, \ldots, S,$$

Together, these two equations imply the following state-space form

$$\widetilde{Z}_{bst} = \begin{bmatrix} \widetilde{\Lambda}_{H.bs0} & \widetilde{\Lambda}_{H.bs1} & \cdots & \widetilde{\Lambda}_{H.bsl_H^*} \end{bmatrix} \begin{pmatrix} H_{bst} \\ H_{bst-1} \\ \vdots \\ H_{bst-l_H^*} \end{pmatrix} + \epsilon_{Zbst}$$

$$\begin{pmatrix} H_{bst} \\ H_{bst-1} \\ \vdots \\ H_{bst-l_H^*} \end{pmatrix} = \begin{pmatrix} \alpha_{G.bst} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{bmatrix} \Psi_{H.bs1} & \cdots & \Psi_{H.bsq_H} & 0 & \cdots & 0 \\ I & 0 & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \vdots & & \vdots \\ 0 & \cdots & I & 0 & \cdots & 0 \end{bmatrix} \begin{pmatrix} H_{bst-1} \\ H_{bst-2} \\ \vdots \\ H_{bst-l_H^*-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{Hbst} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

or

$$\widetilde{Z}_{bst} = \vec{\widetilde{\Lambda}}_{H.bs}\vec{H}_{bst} + \vec{\epsilon}_{Zbst} \tag{14}$$
$$\vec{H}_{bst} = \vec{\alpha}_{G.bst} + \vec{\Psi}_{Hbs}\vec{H}_{bst-1} + \vec{\epsilon}_{Hbst} \tag{15}$$

For blocks that do have a subblock structure, the observation and state equation at the block level become

$$\widetilde{H}_{bt} = \begin{bmatrix} \widetilde{\Lambda}_{G.b0} & \widetilde{\Lambda}_{G.b1} & \cdots & \widetilde{\Lambda}_{G.bl_G^*} \end{bmatrix} \begin{pmatrix} G_{bt} \\ G_{bt-1} \\ \vdots \\ G_{bt-l_G^*} \end{pmatrix} + \epsilon_{Hbt}$$

$$\begin{pmatrix} G_{bt} \\ G_{bt-1} \\ \vdots \\ G_{bt-l_G^*} \end{pmatrix} = \begin{pmatrix} \alpha_{F.bt} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{bmatrix} \Psi_{G.b1} & \cdots & \Psi_{G.bq_G} & 0 & \cdots & 0 \\ I & 0 & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \vdots & & \vdots \\ 0 & \cdots & I & 0 & \cdots & 0 \end{bmatrix} \begin{pmatrix} G_{bt-1} \\ G_{bt-2} \\ \vdots \\ G_{bt-l_G^*-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{Gbt} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

or

$$\widetilde{H}_{bt} = \vec{\widetilde{\Lambda}}_{G.b}\vec{G}_{bt} + \vec{\epsilon}_{Hbt} \tag{16}$$
$$\vec{G}_{bt} = \vec{\alpha}_{F.bt} + \vec{\Psi}_{G.b}\vec{G}_{bt-1} + \vec{\epsilon}_{Gbt} \tag{17}$$

C Tables and Figures

Table 1: Univariate Analysis of Principal Component Estimates
Two Step Model:

$$\begin{aligned}
\widetilde{G}_{bjt} &= \widetilde{\Lambda}_{F.bj}\widetilde{F}_t(\widetilde{G}_t) + \widetilde{e}_{Gbjt} \\
\widetilde{G}_{bjt} &= \widetilde{\Psi}_{G.bj}\widetilde{G}_{bjt-1} + \widetilde{\epsilon}_{Gbjt} \\
\widetilde{e}_{Gbjt} &= \widetilde{\Psi}_{e_{Gbj}}\widetilde{e}_{Gbjt-1} + \widetilde{\epsilon}_{Gbjt} \\
\widetilde{F}_{kt}(\widetilde{G}_t) &= \widetilde{\Psi}_{F.k}\widetilde{F}_{kt-1}(\widetilde{G}_t) + \widetilde{\epsilon}_{Fkt}.
\end{aligned}$$

Block	T	N_b	IC_2	$R^2_{G_{b.1}}$	$\widetilde{\Psi}_{\widetilde{G}.b1}$
CU	207	25	1	0.249	0.376
IP	207	38	2	0.267	0.480
ES	207	82	3	0.205	0.281
HS	207	92	8	0.115	0.131
MS	207	35	4	0.171	0.160
DG	207	60	2	0.151	0.562

Note: Let $\widetilde{G}_{bjt}$ be the j-th factor obtained by the method of principal components using data from block b. $R^2_{G_{b.1}}$ is the explanatory power of the $1st$ factor, obtained as the ratio of $1st$ largest eigenvalue $X'X$ to the sum of the eigenvalues. $\widetilde{\Psi}_{\widetilde{G}.b1}$ is the estimated first order autocorrelation coefficient of $\widetilde{G}.b1$.

Table 2: A Three Level Model for Production with Six Blocks:

$$
\begin{aligned}
X_{bit} &= \Lambda_{G.bi}(L)G_{bt} + e_{Xbit} \\
G_{bt} &= \Lambda_{F.b}(L)F_t + e_{Gbt} \\
\psi_{F.k}(L)F_{kt} &= \epsilon_{Fkt}, \quad k = 1, \dots K_F \\
\psi_{G.bj}(L)e_{Gbjt} &= \epsilon_{Gbjt}, \quad j = 1, \dots K_{Gb} \\
\psi_{X.bi}(L)e_{Xbit} &= \epsilon_{Xbit}, \quad i = 1, \dots N_b.
\end{aligned}
$$

Block	j	$\widehat{\psi}_{G.bj}$	$\widehat{\sigma}^2_{\epsilon_{Gbj}}$	S.E	
CU: 1	1	0.373	0.064	0.113	0.021
CU: 1	2	-0.122	0.057	0.091	0.016
IP: 2	1	0.170	0.015	0.110	0.007
IP: 2	2	-0.140	0.047	0.089	0.017
ES: 3	1	0.052	0.015	0.137	0.007
ES: 3	2	-0.160	0.031	0.115	0.010
HS: 4	1	0.198	0.137	0.095	0.031
HS: 4	2	-0.069	0.056	0.096	0.010
MS: 5	1	0.436	0.824	0.128	0.091
MS: 5	2	0.059	0.111	0.093	0.024
DG: 6	1	-0.013	0.030	0.172	0.007
DG: 6	2	-0.009	0.030	0.175	0.006
Factor		$\widehat{\psi}_{F.k}$	$\widehat{\sigma}^2_{F.k}$	S.E.	
1		0.880	0.061	0.040	0.017

Decomposition of Variance

	Estimates			Standard Errors		
block	share_F	share_G	share_X	share_F	share_G	share_X
1 CU:	0.303	0.144	0.553	0.069	0.021	0.055
2 IP:	0.321	0.131	0.549	0.075	0.021	0.058
3 ES:	0.279	0.114	0.607	0.073	0.020	0.056
4 HS:	0.081	0.150	0.769	0.034	0.013	0.026
5 MS:	0.117	0.222	0.661	0.056	0.033	0.031
6 DG:	0.101	0.123	0.777	0.044	0.014	0.035

Note: This table provides estimates $\widehat{\psi}_{G.bj}$ and $\widehat{\psi}_{F.k}$ governing the serial correlation in the block-specific and common components e_{Gbj} and F_{kt}. It also gives estimates $\widehat{\sigma}^2_{\epsilon_{Gbj}}$ and $\widehat{\sigma}^2_{F.k}$ governing the variance of the innovations ϵ_{Gbjt} and ϵ_{Fkt}. The lower panel documents the shares of variance in X_{bit} due to $\epsilon_{Fkt}, \epsilon_{Gbjt}$, and ϵ_{Xbit} averaged across the series in each block. All reported estimates and standard errors are the posterior means and posterior standard deviations of the draws retained after an initial burn-in.

Table 3: Correlation Between $\widehat{G}_{bjt}$ and $\widetilde{e}_{kt}$

k	b	j	R^2
1	3	2	0.17
2	3	1	0.20
2	3	2	0.32
2	5	1	0.13
3	5	1	0.27
4	4	2	0.16
5	1	2	0.60
5	2	2	0.31
7	6	2	0.15

Note: This table provides estimates of the correlation between block-specific factors $\widehat{G}_{bjt}$ from our three level model of real activity and residuals $\widetilde{e}_{kt}$ obtained by regressing the k-th factor estimated by principal components on our estimate $\widehat{F}_t$. R^2 is the coefficient of determination from regressing $\widetilde{e}_{kt}$ on the j-th factor in block b.

Table 4: A Four Level Model for Real Activity with Five Blocks and Fourteen Subblocks

$$
\begin{aligned}
Z_{bsit} &= \Lambda_{H.bsi}(L)H_{bst} + e_{Zbsit}, \quad \psi_{Z.bsi}(L)e_{Zbsit} = \epsilon_{Zbsit}, \\
H_{bst} &= \Lambda_{G.bs}(L)G_{bt} + e_{Hbst}, \quad \psi_{H.bs}(L)e_{Hbst} = \epsilon_{Hbst}, \\
G_{bt} &= \Lambda_{F.b}(L)F_t + e_{Gbt}, \quad \psi_{G.b}(L)e_{Gbt} = \epsilon_{Gbt}, \\
\psi_F(L)F_t &= \epsilon_{Ft}.
\end{aligned}
$$

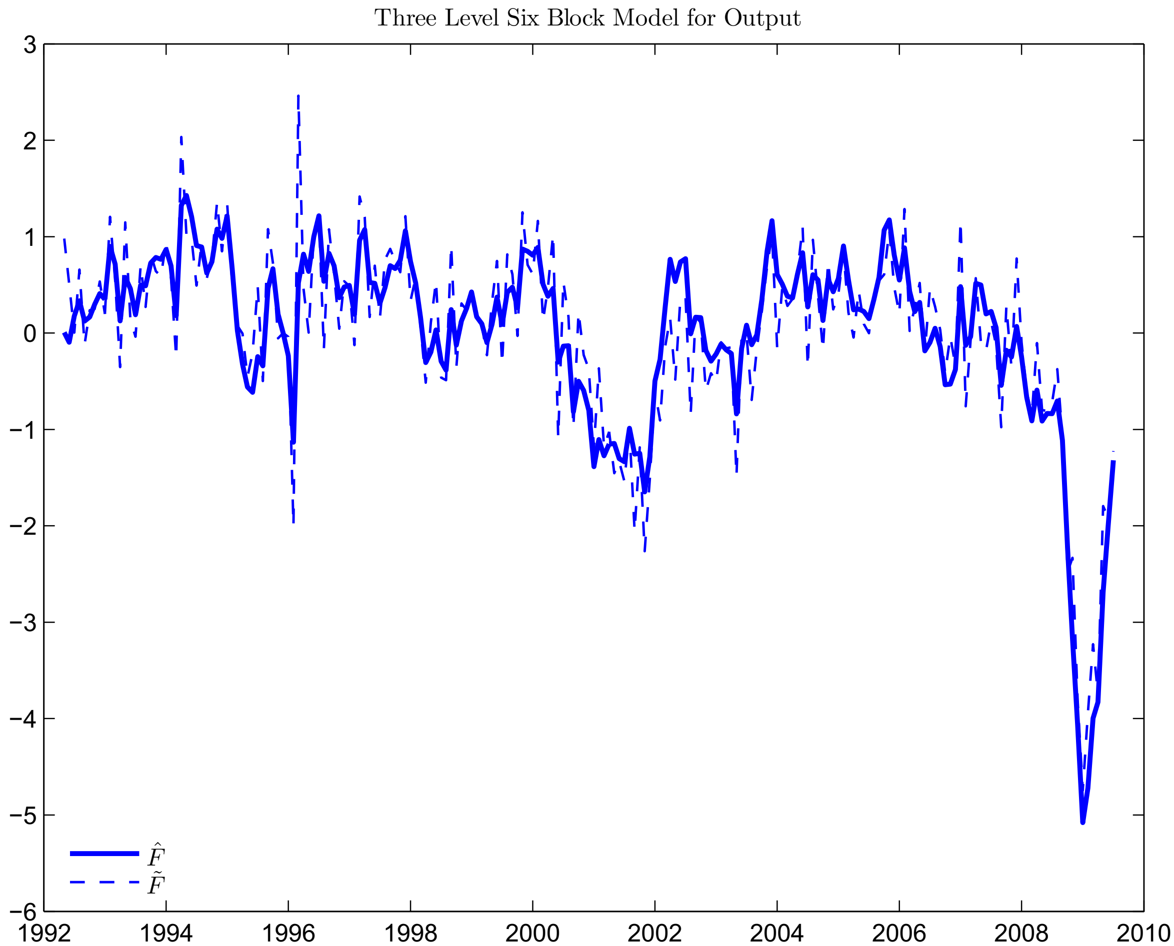
b	$\widehat{\psi}_{G.b}$	$\widehat{\sigma}^2_{\epsilon_{Gb}}$	S.E	
Output	0.396	0.044	0.162	0.016
Employment	0.537	0.019	0.186	0.008
Demand	0.177	0.055	0.161	0.021
Housing	0.706	0.021	0.200	0.010
Mfg Surveys	0.550	0.078	0.143	0.043
Factor	$\widehat{\psi}_F$	$\widehat{\sigma}^2_F$	S.E.	
1	0.824	0.007	0.183	0.004

Decomposition of Variance

block	sub-block	share_F	share_G	share_H	share_X
		Estimates			
Output	IP	0.090	0.115	0.176	0.619
Output	CU	0.089	0.113	0.161	0.637
Output	DG	0.022	0.026	0.180	0.773
Employment	HS	0.042	0.090	0.226	0.642
Employment	ES	0.015	0.033	0.171	0.782
Demand	RS	0.030	0.073	0.212	0.685
Demand	WT	0.012	0.028	0.153	0.806
Demand	AUTO	0.018	0.049	0.393	0.539
Housing	H-starts	0.005	0.069	0.176	0.750
Housing	H-newsales	0.002	0.030	0.285	0.683
Housing	H-existsales	0.003	0.038	0.626	0.333
Mfg Surveys	ISM	0.035	0.161	0.248	0.556
Mfg Surveys	PhilaFed	0.010	0.045	0.197	0.747
Mfg Surveys	ChicagoFed	0.022	0.071	0.461	0.446
		Standard Errors			
Output	IP	0.059	0.030	0.031	0.043
Output	CU	0.060	0.028	0.030	0.042
Output	DG	0.023	0.011	0.019	0.023
Employment	HS	0.037	0.030	0.025	0.033
Employment	ES	0.013	0.019	0.014	0.020
Demand	RS	0.028	0.020	0.025	0.027
Demand	WT	0.013	0.012	0.015	0.017
Demand	AUTO	0.023	0.038	0.049	0.039
Housing	H-starts	0.006	0.042	0.020	0.035
Housing	H-newsales	0.003	0.035	0.033	0.036
Housing	H-existsales	0.004	0.040	0.052	0.041
Mfg Surveys	ISM	0.032	0.065	0.065	0.035
Mfg Surveys	PhilaFed	0.014	0.026	0.026	0.026
Mfg Surveys	ChicagoFed	0.041	0.055	0.072	0.034

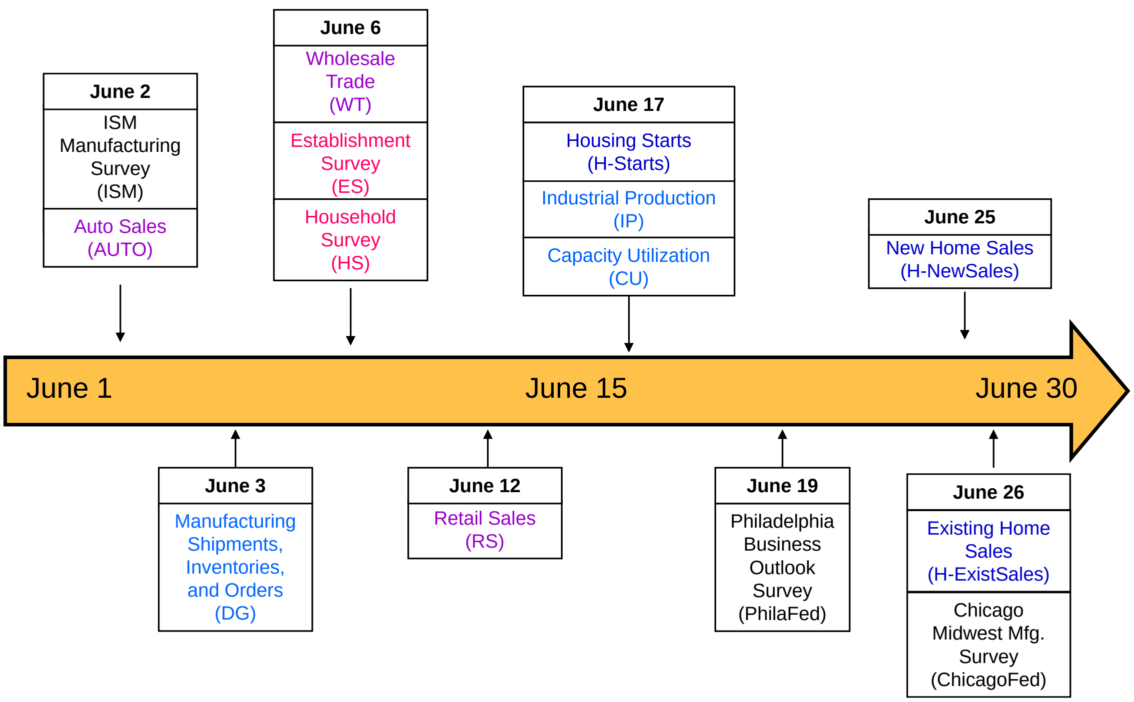
Note: This table provides estimates $\widehat{\psi}_{G.b}$ and $\widehat{\psi}_F$ governing the serial correlation in the block-specific and common components e_{Gbt} and F_t. It also gives estimates $\widehat{\sigma}^2_{\epsilon_{Gb}}$ and $\widehat{\sigma}^2_F$ governing the variance of the innovations ϵ_{Gbt} and ϵ_{Ft}. The lower panel documents the shares of variance in Z_{bsit} due to $\epsilon_{Ft}, \epsilon_{Gbt}, \epsilon_{Hbst}$, and ϵ_{Zbsit} averaged across the series in each subblock. All reported estimates and standard errors are the posterior means and posterior standard deviations of the draws retained after an initial burn-in.

Figure 1: Three Level Model of Real Activity with 6 Blocks



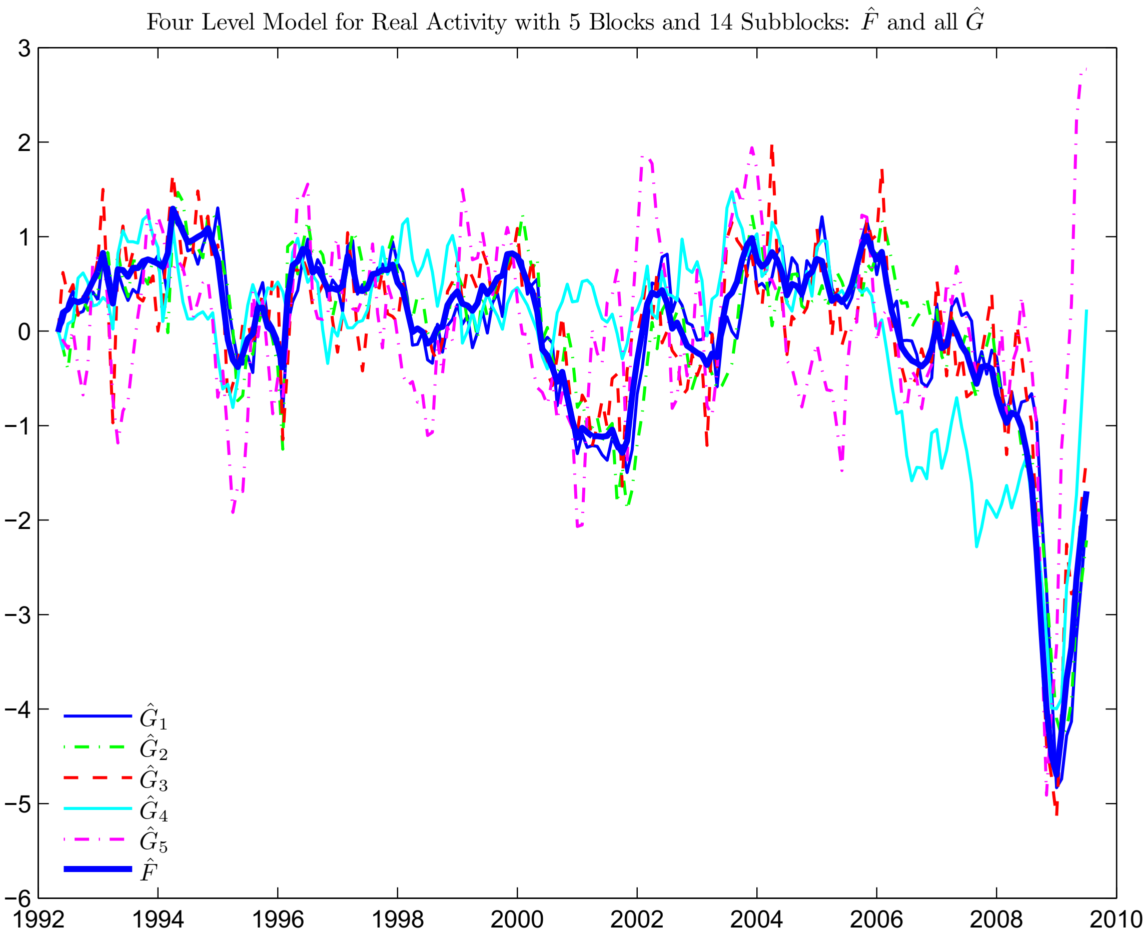
Note: This figure plots the common factor estimate $\widehat{F}$ from our three level model of real activity along with the first principal component $\widetilde{F}$ of the entire data panel.

Figure 2: Release Schedule for Output Related Data in June 2009



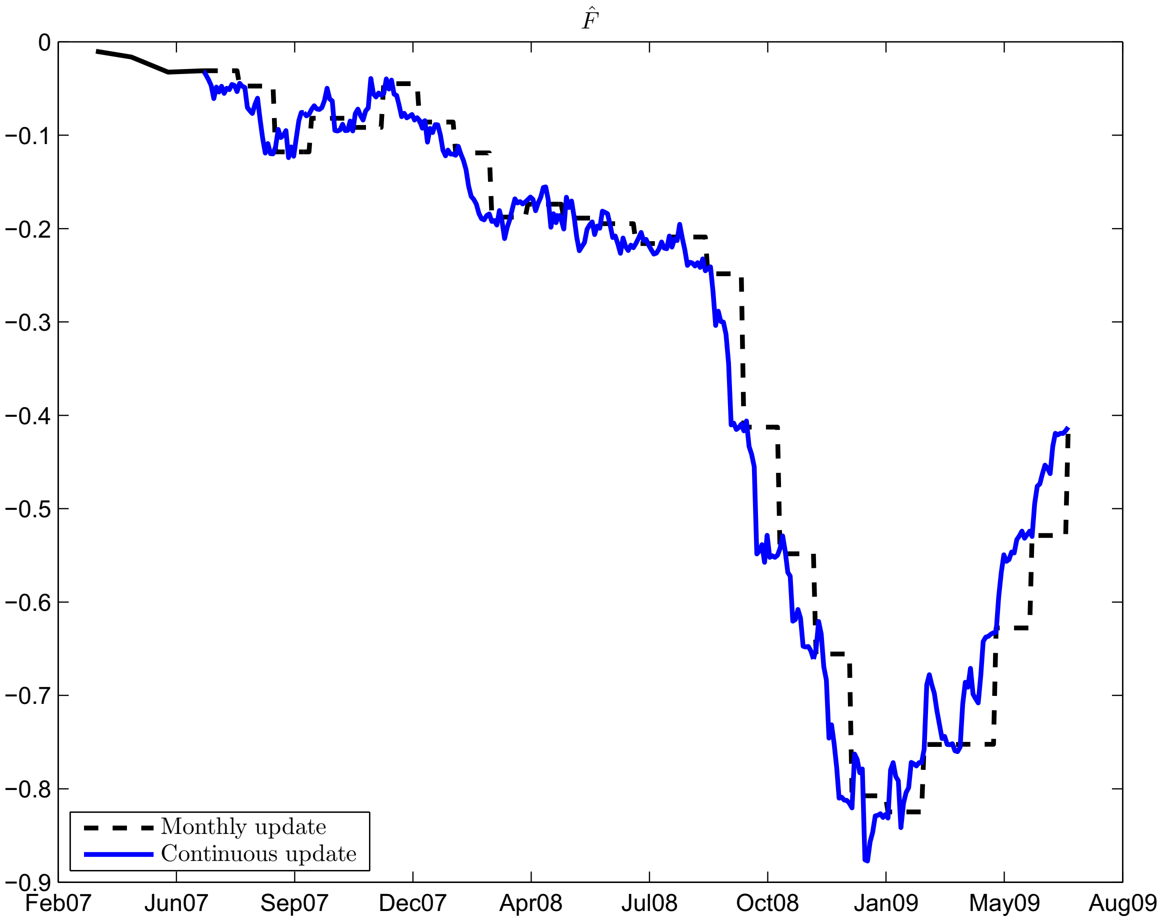
Note: This figure provides a schematic of the timing of the different data releases which constitute the fourteen subblocks of our four level model over the course of June 2009.

Figure 3: Four Level Model of Real Activity with 5 Blocks and 14 Subblocks



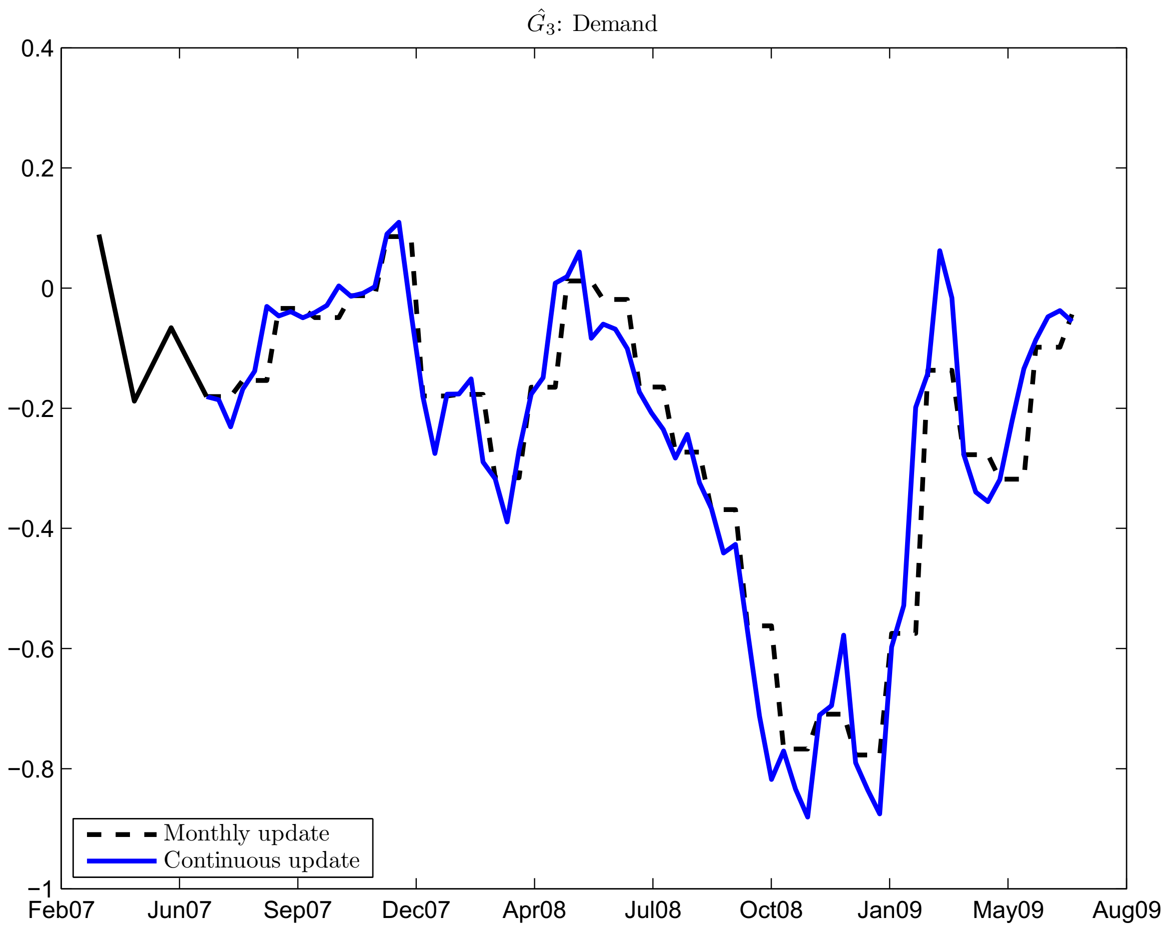
Note: This figure plots the common factor estimate $\widehat{F}$ from our four level model of real activity along with the block-specific factors $\widehat{G}$ for the five blocks of data representing the different categories of real activity-related information that we employ. These are Output ($b = 1$), Demand ($b = 2$), Employment ($b = 3$), Housing ($b = 4$), and Manufacturing Surveys ($b = 5$).

Figure 4: Monitoring Real Activity in a Four Level Model with 5 Blocks and 14 Subblocks



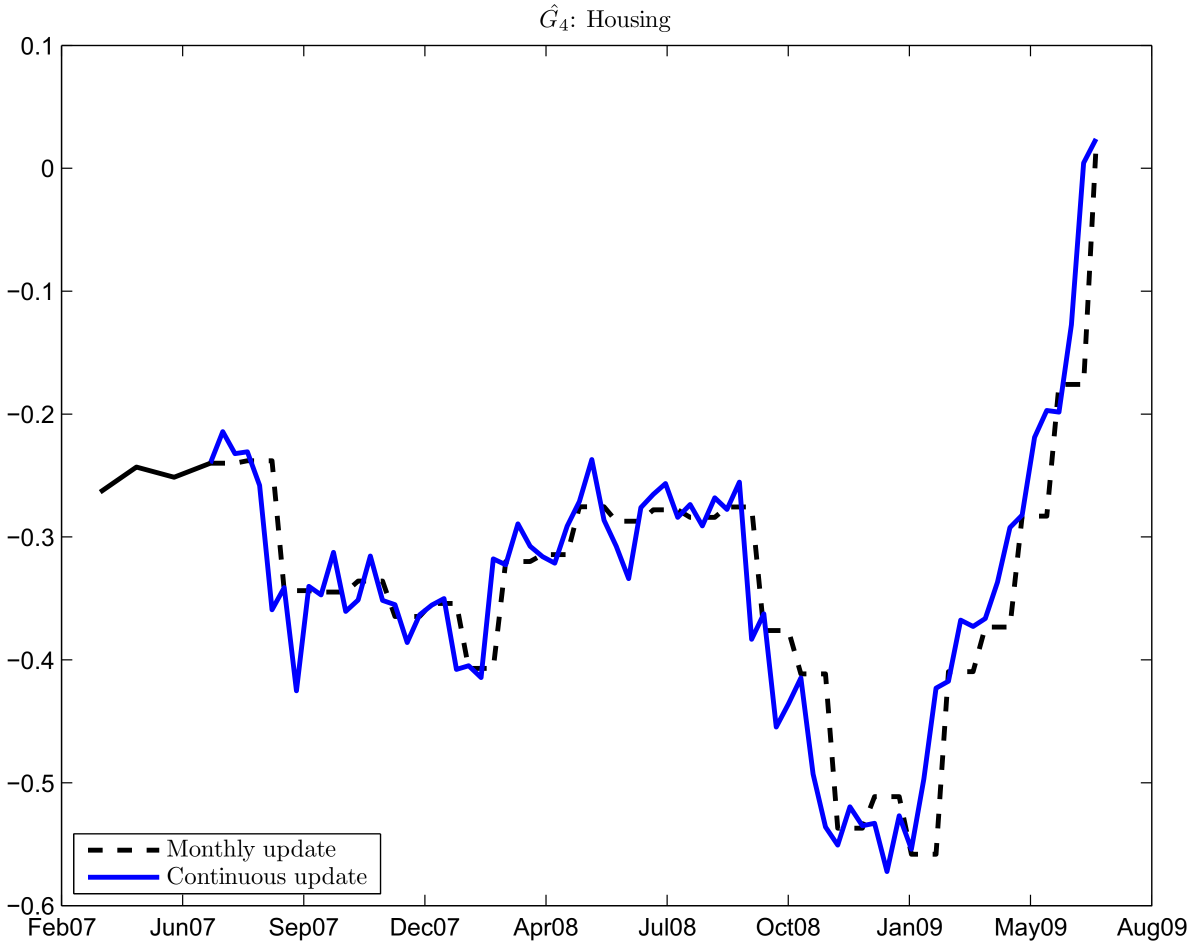
Note: This figure plots the continuous update of the common factor $\widehat{F}$ from our four level model, obtained whenever there is a new release in any of the fourteen subblocks, along with the update of the factor obtained with a full new set of observations at the end of each month.

Figure 5: Monitoring Demand in a Four Level Model with 5 Blocks and 14 Subblocks



Note: This figure plots the continuous update of the block-specific factor $\widehat{G}_b$ for the Demand block from our four level model, obtained whenever there is a new release in any of the three subblocks Retail Sales, Wholesale Trade or Auto Sales, along with the update of the factor obtained with a full new set of observations at the end of each month.

Figure 6: Monitoring Housing Activity in a Four Level Model with 5 Blocks and 14 Subblocks



Note: This figure plots the continuous update of the block-specific factor $\widehat{G}_b$ for the Housing block from our four level model, obtained whenever there is a new release in any of the three subblocks Housing Starts, New Home Sales or Existing Home Sales, along with the update of the factor obtained with a full new set of observations at the end of each month.

Figure 7: Monitoring Manufacturing Activity in a Four Level Model with 5 Blocks and 14 Subblocks

Note: This figure plots the continuous update of the block-specific factor $\widehat{G}_b$ for the Manufacturing Survey block from our four level model, obtained whenever there is a new release in any of the three subblocks ISM, Philadelphia Fed Business Outlook Survey or Chicago Fed MidWest Mfg Survey, along with the update of the factor obtained with a full new set of observations at the end of each month.