Dynamic Hierarchical Factor Models*

Serena Ng[†] Emanuel Moench[‡] Simon Potter[§]

December 11, 2008

Abstract

This paper presents a three level dynamic factor model with a block structure that allows us to distinguish series-specific variations from two types of common variations:- those from (level two) factors that are common to units within a block, and those from (level three) factors that are common across blocks. A four-level model is obtained by allowing each block to have sub-blocks. The model is estimated using an MCMC algorithm that explicitly takes into account the dependence of the lower level factors on the higher level ones. We find that regional variations dominate aggregate level variations in housing prices and that in spite of the attention given to releases of employment data, the aggregate component in the employment block is very small.

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

JEL classification: C10, C20, C30

*We would like to thank Evan LeFlore for excellent research assistance on this project. The first 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.

[†]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. emanuel.moench@ny.frb.org

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

1 Introduction

Macroeconomists define the business cycle as common fluctuations across a wide range of economic variables. Starting with the seminal work by Geweke (1977), dynamic factor models have been used to extract the common fluctuations in economic variables. As predicted by economic theory and verified by numerous empirical papers with Sargent and Sims (1977) being the first, business cycle fluctuations can be captured by a relatively small number of factors. However, the number of observed time series relevant for business cycle analysis is large. This has led to intensive research to estimate common factors from large panels of data under weak conditions, see Stock and Watson (2002) and Forni et al. (2005) among others. Thus, while early work extracts common factors from a relatively small number of series, the more modern practice is to extract them from a large number of time series chosen to give a balanced representation to the various sectors of the economy.

In this paper, we develop a state space estimation methodology for a highly flexible hierarchical (multi-level) dynamic factor model. The hierarchy is produced by splitting a large panel of data into a much smaller number of blocks, each of which consists of a reasonably large number of series. We assume that each block is driven by factors that are common amongst series within the block, and factors at the block level are themselves driven by factors that are common to all block-level factors. Additional levels can be obtained by further splitting blocks into finer sub-blocks and assuming that factors at the sub-block level are driven by a smaller number of block-level factors. This hierarchical structure implies that the transition equation for the common factors at a given level has a time varying intercept that depends on the common factors at the next higher level and must be taken into account in the filtering algorithm.

The point of departure of our analysis is to organize a large panel of data into smaller blocks, and there are several motivations for doing so. Data on the US economy are published on a continuous schedule. For example, in the first week of each month the Bureau of Labor Statistics (BLS) publishes the employment report which consists of an establishment and a household survey. The release of the employment report is followed by the release of retail sales, retail inventories and inventory-sales data by the Commerce Department in the second week of each month. This is followed by the release of industrial production and capacity utilization data by the Board of Governors in the third week of each month and so on. Organizing the data into an employment block, a retail sales block, and so forth has two advantages. First, exploiting the sequence in which data arrive is useful for real-time monitoring of economic activity. Second, we can explicitly distinguish between comovement at the block level from those at the aggregate level, which is useful for understanding the observed fluctuations in the data.

Sectoral and geographic structure provide another natural way to organize the data into blocks. For example, housing data are available for different regions in the U.S. With each region corresponding to a block, we can distinguish between housing shocks at the regional and the national levels. To facilitate economic analysis, it is often convenient to consider a price factor, a financial market factor, and a labor market factor. This can be achieved by estimating the factors from a block of price data, a block of labor market data, a block of financial time series and so on.[1] Since the blocks have a well-defined interpretation, one can give names to the block-level factors immediately. This overcomes a common criticism of large dimensional factor analysis that the factors are difficult to interpret.

To provide a simple example of a hierarchical dynamic factor model, let $t = 1, \ldots, T$ be an index for time, and $b = 1, \ldots, B$ be the index for blocks. Abstracting from dynamics to focus on the hierarchical structure for now, an observation on variable i belonging to block b observed at time t, denoted x_{bit}, is modeled as

$$x_{bit} = \lambda'_{G.bi} G_{bt} + e_{Xbit}, \qquad i = 1, ..., N_b \tag{1}$$

where G_{bt} is a set of k_{Gb} latent block-level factors, N_b is the total number of variables in block b and e_{Xbit} is a purely idiosyncratic error (i.e., independent of all other errors and factors but not necessarily an IID process). We posit that the latent block-level factors themselves obey a factor structure, viz:

$$G_{bkt} = \lambda'_{F.bk} F_t + e_{Gbkt} \tag{2}$$

where F_t is a vector of k_F factors that are common across blocks and e_{Gbkt} is a block-specific idiosyncratic error that is independent of all other block specific errors and all factors but is not necessarily an IID process. Thus, variables within a block can be correlated through F_t or e_{Gbkt}, but variables between blocks can be correlated only through F_t. In the terminology of multilevel models, (1) is the level-1 equation, and (2) is the level-2 equation. A stochastic process for F_t would

[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 one of our main applications data are placed into blocks by data source. For example, we might have a retail sales block based on the underlying detail of the the Census Bureau's monthly retail sales release.

constitute a level-3 equation. To be concrete, consider a level-3 model using four blocks of housing data: the Northwest, the West, the Midwest, and the South. For each block (or region), we have data (x_{bit}) on house prices, houses sold, building permits, etc. Then G_{2t} is the set of (block-level) factors common to variables in the West, and e_{G2t} are shocks to the West. Idiosyncratic shocks to each series collected for the West are represented by e_{X2it}. The factors F_t represent variations common to all four regions and are the only source of co-movement between, for example, the West and the Northeast.

The difference between a multilevel and a two-level factor model is best understood when F_t and G_{bt} are scalars. With $k_G = K_F = 1$,

$$
\begin{aligned}
x_{bit} &= \lambda_{G.bi}(\lambda_{F.b1}F_t + e_{Gb1t}) + e_{Xbit} \\
&= \lambda_{bi}F_t + v_{bit} \qquad (3)
\end{aligned}
$$

where $\lambda_{bi} = \lambda_{G.bi}\lambda_{F.b1}$ and

$$v_{bit} = \lambda_{G.bi}e_{Gb1t} + e_{Xbit}.$$

A standard factor model ignores the block structure and stacks all observations up irrespective of which block an observation belongs to. The data are thus analyzed using the level-2 representation

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

This level-2 representation would be equivalent to an exact factor model if $\{e_{Gb1t} : b = 1, ..., B\}$ was a zero stochastic process. 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 making this hard to impose assumption, our hierarchical model tackles this problem by explicitly specifying the block structure.

Modeling the dynamics at the individual, block, and the aggregate levels is useful in a variety of economic models. For example, risk factors in asset returns can be due to variations on the market portfolio, which is common, and to industry specific effects. It is also reasonable to assume that business cycle variations have a global, a regional or a country-specific component. Our hierarchical dynamic factor model naturally lends itself to analyzing such empirical questions.

This paper sets up the model and the framework for estimation. Our goal here is to show that common variations at the block level are significant, and that going from a two to a three or higher level model not only has statistical advantages, it also provides a better understanding of economic fluctuations. We consider three examples. The first uses housing data where the blocks are determined by geography. The second is a three-level model of production with six blocks of data organized by the timing of data releases. The third is a four-level model for real economic activity. The decomposition of variance in each case shows that block-level variations tend to be stronger than the common variations, though both are small relative to the purely idiosyncratic components in the series. We also compare the factor estimates to the principal component estimates and isolate one episode in which the principal components estimator treats a block-level variation as common because the large magnitude of the block-level event contributes significantly to the total variations in the data.

2 A Three Level 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_b common factors G_b in each block $b = 1, \ldots, B$. Let the mean zero block-level factors be G_{bkt} for $k = 1, \ldots k_b$. Hence, there is a total number $K_G = (k_1 + \ldots + k_B)$ 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 in a given block b is decomposed into a serially correlated idiosyncratic component, e_{Xbit}, and a common component $\Lambda_{G.bi}(L)G_{bt}$ which it shares with other variables in the same block. Each block-level factor G_{bt} has a serially correlated block-specific component e_{Gbt} and a common component $\Lambda_{F.b}(L)F_t$ which it shares with all other blocks. Finally, the economy-wide factors F_t are assumed to be serially correlated. Let

$$
\begin{aligned}
X_{bt} &= \begin{pmatrix} X_{bt.1} & X_{bt.2} & \ldots X_{bt.N_b} \end{pmatrix}' \\
G_{bt} &= \begin{pmatrix} G_{bt.1} & G_{bt.2} & \ldots G_{bt.k_b} \end{pmatrix}'.
\end{aligned}
$$

Then, the model can be summarized by the following equations:

$$X_{bt} = \Lambda_{G.b0}G_{bt} + \ldots + \Lambda_{G.bs_{Gb}}G_{b,t-s_{Gb}} + e_{Xbt}, \tag{4}$$

$$G_{bt} = \Lambda_{F.b0}F_t + \ldots + \Lambda_{F.bs_F}F_{t-s_F} + e_{Gbt} \tag{5}$$

$$F_t = \Psi_{F.1}F_{t-1} + \ldots + \Psi_{F.q_F}F_{t-q_F} + \epsilon_{Ft} \tag{6}$$

$$e_{Xbit} = \Psi_{X.bi1}e_{Xbi,t-1} + \ldots + \Psi_{X.biq_{Xbi}}e_{Xbi,t-q_{Xbi}} + \epsilon_{Xbit} \tag{7}$$

$$e_{Gbt} = \Psi_{G.b1}e_{Gb,t-1} + \ldots + \Psi_{G.bq_{Gb}}e_{Gb,t-q_{Gb}} + \epsilon_{Gbt}. \tag{8}$$

The idiosyncratic components e_{Xbi} are AR processes of order q_{Xbi}, the block-specific component is an AR process of order q_{G_b}, and the economy-wide factors F_{kt} are AR processes of order q_{F_k}. We assume normally distributed innovations throughout. Thus,

$$\begin{aligned}
\epsilon_{Xbi} &\sim N(0, \sigma^2_{Xbi}) \quad i = 1, \ldots N_b \\
\epsilon_{Gbk} &\sim N(0, \sigma^2_{Gbk}) \quad k = 1, \ldots k_{Gb}, b = 1, \ldots B \\
\epsilon_{F_r} &\sim N(0, \sigma^2_{F_r}) \quad r = 1, \ldots K_F.
\end{aligned}$$

The dynamics of the model can be enriched by allowing for stochastic volatility and Markov switching effects at different levels of the hierarchy. The current specification allows the lag order of the factor loading matrix 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. Thus, $s_{Gb} = (s_{Gb.1}, \ldots s_{Gb.N_b})$ is a vector. Similarly, $q_{Xb} = (q_{Xb.1}, \ldots q_{Xb.N_b})$ and $q_{Gb} = (q_{Gb.1}, \ldots q_{Gb.k_b})$ are also vectors with possibly non-identical entries. Let $s_F = \max_{b \in B} s_{Fb}$, $s_G = \max_b(\max_{i \in N_b} s_{Gb.i})$, $q_X = \max_b(\max_{i \in N_b} q_{Xbi})$, and $q_G = \max_b(\max_{k \in k_b} q_{Gb.k})$. Stacking up the data by blocks and letting

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

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 s_G, $\Lambda_F(L)$ is a $K_G \times K_F$ matrix polynomial of order s_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 $s = 0, \ldots, s_G$

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

The block-diagonal structure of $\Lambda_{G.s}$ implies that each block X_b of variables exclusively loads on the block-level factors G_b. We assume that at lag 0, and for each $b = 1, \ldots, B$, $\Lambda_{G.b0}$ is a $N_b \times k_b$ matrix whose upper-left $k_b \times k_b$ block is lower-triangular with ones on the diagonal. For example, if $k_b = 2$, 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 implies that the first variable within each block exclusively loads on the contemporaneous observation of the first block-level factor, the second exclusively on the contemporaneous observations of the first two block-level factors and so on. We order the variables in each block such that the first k_b series are variables of economic interest and yet have independent information. The loadings of the remaining variables in each block are unrestricted.

To ensure identification of the economy-wide factors F, we assume that the upper $K_F \times K_F$ submatrix of $\Lambda_{F.0}$ is lower triangular. If $K_F = 2$, we have

$$\Lambda_{F.0} = \begin{bmatrix} 1 & 0 \\ \Lambda_{F.0_{2,1}} & 1 \\ \Lambda_{F.0_{3,1}} & \Lambda_{F.0_{3,2}} \\ . & . \\ \Lambda_{F.0_{K_G,1}} & \Lambda_{F.0_{K_G,2}} \end{bmatrix}$$

This normalization assumes that the blocks are ordered so that the common factors load heavily on the first block-level factors. An alternative is to normalize the variance of ϵ_F to unity and to restrict the diagonal elements of the upper-left $K_F \times K_F$ block of $\Lambda_{F.0}$ to be positive.

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. 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

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

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, $f(x_t; G_t; \Theta_G)$ is the measurement equation and $f(G_t|F_t; \Theta_F)$ 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 $\Psi_F(L) = I_{K_F} - \Psi_{F.1} L - \dots \Psi_{F.q_F} L^{q_F}$. This can be rewritten 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

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

where $\epsilon_{Ft} \sim N(0, \Sigma_F)$ and $||\vec{\Psi}_F|| < 1$.

Block-Level Dynamics The structural (pseudo measurement) equation is

$$\begin{aligned} G_t &= \Lambda_{F.0} F_t + \Lambda_{F.1} F_{t-1} + \ldots + \Lambda_{F.s_F} F_{t-s_F} + e_{Gt} \\ &= \Lambda_F(L) F_t + e_{Gt} \end{aligned} \tag{9}$$

with $G_t' = \begin{pmatrix} G_{1t} & \vdots & G_{Bt} \end{pmatrix}$, G_{bt} being a $k_b \times 1$ vector,

$$\Psi_G(L) e_{Gt} = \varepsilon_{Gt} \tag{10}$$

and where $\Lambda_{F.s}$ is a $K_G \times K_F$ matrix of factor loadings. We call this a structural instead of a measurement equation because G_t is not observed. Here, $\Psi_G(L)$ is a block-diagonal matrix with diagonal blocks $\Psi_{Gb}(L)$, where for $b = 1, \ldots B$, $\Psi_{G.b}(L)$ is itself a diagonal matrix with elements $\psi_{G.bi}(L)$,

$$\psi_{G.bi}(L) = 1 - \psi_{G.bi1} L - \ldots - \psi_{G.biq_G} L^{q_G}$$

We restrict $||\psi_{G.bi}(L)|| < 1$ for stationarity and assume $\epsilon_{Gt} \sim N(0, \Sigma_G)$. Together, (9) and (10) 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

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

where

$$\alpha_{Ft} = \Psi_G(L) \Lambda_F(L) F_t. \tag{11}$$

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

$$\begin{aligned} X_{bt} &= \Lambda_{G.b0} G_{bt} + \ldots + \Lambda_{G.bs_G} G_{bt-s_G} + e_{Xbt}, \\ \Psi_{X.b}(L) e_{Xbt} &= \varepsilon_{Xbt} \end{aligned}$$

where we recall that the top $k_{Gb} \times k_{Gb}$ block of $\Lambda_{G.b0}$ is lower triangular with diagonal elements equal to one, and with X_{bt} ordered so that G_{bt} have non-trivial loadings on the first k_b variables in block b. For each b, $\Psi_{X.b}(L)$ is a diagonal matrix with elements given by

$$\psi_{X.bi}(L) = 1 - \psi_{X.bi1} L - \ldots - \psi_{X.biq_X} L^{qX}$$

Then, for $b = 1, \ldots B$, the measurement equation for each block can be rewritten as

$$\Psi_{X.b}(L) X_{bt} = \Psi_{X.b}(L) \Lambda_{G.b}(L) G_{bt} + \varepsilon_{Xbt},$$

or

$$\widetilde{X}_{bt} = \widetilde{\Lambda}_{G.b}(L) G_{bt} + \varepsilon_{Xbt} \tag{12}$$

where $\widetilde{X}_{bt} = \Psi_{X.b}(L) X_{bt}$ and $\widetilde{\Lambda}_{G.b}(L) = \Psi_{X.b}(L) \Lambda_{G.b}(L)$. These can be stacked to produce

$$\widetilde{X}_t = \widetilde{\Lambda}_G(L) G_t + \varepsilon_{Xt}$$

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

$$vec(Var(X_{bi})) = \gamma'_{F.bi} vec(Var(F_t)) + \gamma'_{G.bi} vec(Var(G_{bt})) + vec(Var(e_{Xbi})) \tag{13}$$

where

$$\begin{aligned} \gamma'_{F.bi} &= \left(\sum_{s=0}^{s_G} \lambda'_{G.bis} \otimes \lambda'_{G.bis} \right) \cdot \left(\sum_{s=0}^{s_F} \Lambda_{F.bs} \otimes \Lambda_{F.bs} \right) \\ \gamma'_{G.bi} &= \left(\sum_{s=0}^{s_G} \lambda'_{G.bis} \otimes \lambda'_{G.bis} \right) \\ vec(Var(F_t)) &= \left[I - \sum_{q=1}^{q_F} (\Psi_{F.q} \otimes \Psi_{F.q}) \right]^{-1} vec(\Sigma_F) \\ vec(var(G_{bt})) &= \left[I - \sum_{q=1}^{q_G} (\Psi_{G.bq} \otimes \Psi_{G.bq}) \right]^{-1} \cdot vec(\Sigma_{G_b}) \\ vec(Var(e_{Xbi})) &= \left[1 - \sum_{q=1}^{q_X} \psi^2_{X.biq} \right]^{-1} \cdot \sigma^2_{Xbi}. \end{aligned}$$

The total variance is the sum of the unconditional variance of the components multiplied by the effective loadings on the components. Dividing the three components on the right hand side of (13) 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$_X$. 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 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 Fruhwirth-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. Details of Steps 2 and 3 are given in the Appendix.

3 Related Work

Multilevel factor models have been considered extensively in the psychology literature. See, for example, Goldstein and Browne (2002). 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. Howover, these models do not allow for dynamics. Dynamic hierarchical linear models were considered by Gammerman and Mignon (1993), but there are no latent variables. Our three-level factor model shares common features with a few approaches that have previously been suggested in the macroeconomic literature. Kose et al. (2003) and Kose et al. (2008) used multi-level factor models to study international business cycle comovements. In their model, 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

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

where F_t is a world factor, G_{bt} is a factor specific to region b, and where e_{bit} is a component specific to variable i in country b.[2]

Our model hierarchical multilevel model differs from theirs in a number of ways. While their F_t and G_{bt} are scalars, we allow for multiple common and multiple block-level factors. Comparing this setup to (3), their loading on the world factor c_i plays the role of our $\lambda_{G.bi}\lambda_{F.b1}$ 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. (2008). In

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

other words, our multilevel model is hierarchical. By imposing this 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. (2008) 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, an issue that has computational consequences as we discuss below.

Diebold et al. (2008) discuss another model that is similar in spirit to ours. These authors decompose interest rates in different countries into global and country-specific level and slope factors. These factors are identified by assuming a particular parametric form of the factor loadings. Diebold et al. (2008) estimate their model in two-steps. First, they obtain individual country factors by separately estimating the level-2 model for each country. Second, they assume a linear dependence of country-level factors on the global factors and obtain the global yield factors by estimating another level-2 model treating the country-level factors as data. In contrast to this two-step approach, we estimate the factors at both levels of the hierarchy simultaneously. In other words, we explicitly take into account the common variation at the third level when estimating the dynamics in the level-2 factors.

Also somewhat related to our model, 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. While we assume the block structure of the data as known, they develop methods to estimate the common, block-level, and idiosyncratic factors from the data. 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 close to 200, and we anticipate using as many as 1,000 series in a full 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 avoid 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.

Since the true F_t are latent variables it will be difficult to compare the precision of the estimates. However, 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). When comparing the results, 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.

4 Empirical Analysis

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.[3] 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.

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. 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.

In the three applications considered in this paper, we use a balanced panel of monthly data from 1992:01-2008:02. After the data transformation, our sample effectively starts in 1992:4, giving $T =$ 191 observations for all blocks. The data are transformed (by taking logarithm and differencing) using Stock and Watson (2008b) as a guide.

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}. The choice of these variables, as well as summary statistics based upon the factors estimated by principal components are reported in Tables A1 and A2.

4.1 A Housing Model Using Actual and Simulated Data

Our first model uses monthly U.S. housing data with a breakdown into three geographical blocks: the Northeast (NE), the West (W), and a third block (CTL) that combines the South with the Midwest. The three blocks comprise 7, 8, and 18 series, respectively. Note that the principal component estimates of the block-level factors will not be precise because of the small number of units in each block. We assume that there are two factors in each block, i.e. $k_{Gb} = 2$ for all b, and

[3] If θ is distributed as inverse χ^2 with ν degrees of freedom and a sale 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.

all block-level factors are in turn driven by one common factor, i.e. $K_F = 1$. The common and block-level factors are modeled to have AR(1) dynamics, that is $q_{Gb} = q_F = 1$. We allow for two lags of the respective factors in the measurement equations, i.e. $s_{Gb} = s_F = 2$. The results are reported in Table 1.

The estimated ψ_F suggest that the common factor is highly persistent, while $\widehat{\psi}_G$ are generally not significantly different from zero. The bottom panel of Table 1 provides the decomposition of variance into economy-wide, regional, and idiosyncratic shocks. At each draw of the Gibbs sampler, we obtain the shares for each series and then average over units within a block. We then report the mean and standard deviation of these block-level decompositions across all draws. According to the results, the common housing factor accounts for 10 to 17% of the regional fluctuations in the housing market while regional factors account for anther 12 to 24%. However, by far the largest source of variations in the housing market are idiosyncratic. Our methodology thus seems capable of disentangling the different levels of cross-sectional variation.

In addition to the housing analysis being of interest in its own right, we also use this simple example to assess the precision of the estimator. Treating the posterior means of the parameter estimates as well as $\widehat{F}_t$ and $\widehat{G}_t$ as 'true' values, and resampling ϵ_X we construct a set of simulated data. The simulated data are then used to estimate the parameters. Comparing the estimates obtained from the simulated data with the 'true' values gives an assessment of the precision of the estimates.

The top left panel of Figure 1 graphs the (true) F against $\widehat{F}$. The two are almost indistinguishable, showing the sampler gives posterior means that are very close to the true values of the latent process. Clearly, $\widehat{F}$ is significantly below zero in recent years. At the end of our estimation sample (2008:2), $\widehat{F}$ is -0.843, with a sample standard deviation of 0.320. In comparison, the maximum value of $\widehat{F}$ over the sample is 0.427. The top right panel of Figure 1 graphs G_{b1} against $\widehat{G}_{b1}$ for the West, which is far more volatile than the factors in the Northeast and the other regions. The $\widehat{G}_b$ are quite close to the true factor processes as implied by the simulated data. It is evident that shocks to all three regions have mostly been negative since 2006. While the state of housing in the other regions appears to be still in a downward trend, there are signs of rebounding in both the West and Northeast at the end of our sample period. In sum, we find that regional and to a smaller extent economy-wide shocks have contributed to the recent housing slump.

4.2 A Six Block Model for Production

We estimate a 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). 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.[4] Our model is described by the following set of parameters : $K_F = 1$, $k_{Gb} = 2$ for all b, $s_{Fr} = 2$, $s_{Gb} = 2$, $q_{Fr} = q_{Gb.k} = q_{Xb.i} = 1$ for all $b = 1, \dots B$, $r = 1, \dots K_F$, $k = 1, \dots K_{Gb}$, 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 .795. 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 $2 \times N$ loadings on G_t, and $K_G \times 1$ loadings on F_t, where $K_G = 12$ and N=315. 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. According to the model, the DG block has the best fit. 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 20% or more of the variation in the data of the block. The block-level shocks roughly explain another 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 the data.

While aggregate shocks to the CU, IP, and ES blocks are more important than the block-level

[4] 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.

shocks, the block-level component is larger than the common component in all remaining blocks. Shocks common to MS block account for around 24% of the variations, compared to the common component of about 4%. The result that stands out in Table 2 is that the idiosyncratic component always explains the largest share of variation. In particular, 80% of the variation in the Household Survey block is idiosyncratic, and only 2% of the variation in that block is explained by the common factor F. Although the monthly employment report (which contains the HS data) is well-watched by financial markets, our findings suggest that the HS data contain little information about the level of non-housing real economic activity. The results generally 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.

The relative importance of the common factors based on principal components estimation is also reported in Table 3. Both one and two step principal component estimation of F_t suggests that the first two factors explain about 40% of the variation in the data. The correlation between our first factor $\widehat{F}_1$ and the first principal component is 0.80.

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 $\widetilde{F}_{rt}$ on $\widehat{F}$ to obtain residuals $\widetilde{e}_{rt}$ for each $r = 1, \ldots K_F$. 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 $\widetilde{e}_{rt}$ on $\widehat{G}_{bkt}$. To conserve space, Table 3 reports the R^2s that exceed 0.1. 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 0.716, while the third principal component is highly correlated with the Household Survey block ($b = 4$). This could be a consequence of the fact that the employment block constitutes 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 Durable Goods block ($b = 6$) are correlated with the second and fourth 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 2 graphs the factors estimated using the three different methods. The top panel plots $\widehat{F}_t$ estimated using our hierarchical model against the first principal component $\widetilde{F}_t(\widetilde{G}_t)$ extracted from the block-level principal components. The lower panel graphs $\widehat{F}_t$ against the first principal component $\widetilde{F}_t$ extracted from the entire data panel. All estimates indicate that the trough of the last recession occurred towards the end of 2001. This is in agreement with the NBER business cycle dates which report November 2001 as the trough of the last recession. All estimates also indicate a slowdown in the level of real activity since the middle of 2005, with the common factor $\widehat{F}_t$ estimated using our hierarchical model suggesting a weaker economy than the principal component estimates.

4.3 A Four Level Model

Some blocks of data are naturally organized by sub-blocks. For example, data on the establishment and household surveys are released together in the employment report, while data for industrial production and capacity utilization are also released at the same time. Our hierarchical factor model can easily be extended to allow for a sub-block 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 sub-blocks from those that do not, let Z_{bsit} be the observed data for block b where s is an index for the sub-blocks. Let H_{bst} be the factors of sub-block b. Then a four-level model can be represented by

$$
\begin{aligned}
Z_{bsit} &= \Lambda_{H.bsi}(L)H_{bst} + e_{Xbsit} &\quad (14)\\
H_{bst} &= \Lambda_{G.bs}(L)G_{bt} + e_{Hbst} &\quad (15)\\
G_{bt} &= \Lambda_{F.b}(L)F_t + e_{Gbt} &\quad (16)\\
F_{kt} &= \Psi_{F.1}F_{k,t-1} + \dots \Psi_{F.q_{Fk}}F_{k,t-q_F} + \epsilon_{Fkt} &\quad (17)\\
e_{Gbt} &= \Psi_{G.b1}e_{Gb,t-1} + \dots + \Psi_{G.bq_{Gb}}e_{Gb,t-q_{Gb}} + \epsilon_{Gbt} &\quad (18)\\
e_{Hbst} &= \Psi_{H.bs1}e_{Hbs,t-1} + \dots + \Psi_{H.bsq_{Hbs}}e_{Hbs,t-q_{Hbs}} + \epsilon_{Hbst} &\quad (19)\\
e_{Xbsit} &= \Psi_{X.bsi1}e_{Xbsi,t-1} + \dots + \Psi_{X.bsiq_{Xbs}}e_{Xbsi,t-q_{Xbs}} + \epsilon_{Xbsit} &\quad (20)
\end{aligned}
$$

The dependence of H_t on G_t implies that

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

where $\alpha_{G.bst} = \Psi_{H.bs}(L)\Lambda_{G.b}(L)G_{bt}$. As in the level three 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 of F_t, we can draw G_{bt} for each b. Blocks that have a sub-block 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 sub-block level, and a multi-level model would be more time intensive to estimate. But conceptually, a model with 'branches' in some but not all blocks is straightforward to set up in our framework.

In our application of a four level model, we consider nine sub-blocks of data: the establishment survey (ES), household survey (HS), manufacturer's survey (MS), durable goods (DG), industrial production (IP) and capacity utilization (CU), retail sales (RS), wholesale trade (WT), and autosales (AUTO). We organize the nine sub-blocks in three blocks that together comprise 402 series. The first is an output block with sub-blocks IP, CU, MS, and DG representing the goods production. The second is a labor market block consisting of sub-blocks ES and HS. The third is a demand block consisting of sub-blocks RS, WT, and AUTO. We estimate one common factor and let $K_G = (1, 2, 1)$ and $K_H = 2$ for all sub-blocks. We interpret the estimated common factor as a factor for real economic activity.

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, while that for the employment block is higher than that 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 sub-block level. The CU and IP blocks continue to have the largest common component. The sub-blocks of the demand block have relatively large variations due to factors common to series in the sub-blocks, but the overall picture remains that idiosyncratic shocks dominate.

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 line is the principal component estimate, $\widetilde{F}_t$, both standardized to have a mean of zero and unit variance. Note that our $\widehat{F}_t$ is noticeably smoother than $\widetilde{F}_t$. The latter features large spikes in 1996 that are also picked up by our sub-block factors for the establishment survey block ES. 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 entire panel puts a lot of weight on this block-level event. In contrast, it is appropriately treated as variations associated to a sub-block of employment using our four-level hierarchical model. Our (non-standardized) estimates suggest that the state of real economic activity at the end of our sample in 2008:02 stood at -.328. With the sample standard deviation of $\widehat{F}_t$ being .202, the level of real activity was thus considerably below average. However, according to our estimate, activity in 2008:02 was still stronger than at the trough of the 2001 recession for which we record a value of -0.598.

5 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.

References

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

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

——— (2006): "Confidence Intervals for Diffusion Index Forecasts and Inference with Factor-Augmented Regressions," *Econometrica*, 74:4, 1133–1150.

Beltratti, A. and C. Morana (2008): "International Shocks and National House Prices," Bocconi University.

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

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

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005): "The Generalized Dynamic Factor Model, One Sided Estimation and Forecasting," *Journal of the American Statistical Association*, 100, 830–840.

Fruhwirth-Schnatter, S. (1994): "Data Augmentation and Dynamic Linear Models," *Journal of Time Series Analysis*, 15, 183–202.

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

Geweke, J. (1977): "The Dynamic Factor Analysis of Economic Time Series," in *Latent Variables in Socio Economic Models*, ed. by D. J. Aigner and A. S. Goldberger, Amsterdam: North Holland.

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

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

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

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

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

——— (2008): "Understanding the Evolution of World Business CYcles," *International Economic Review*, 75, 110–130.

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

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

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

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

Sargent, T. and C. Sims (1977): "Business Cycle Modelling without Pretending to have too much a Priori Economic Theory," in *New Methods in Business Cycle Research*, ed. by C. Sims, Minneapolis: Federal Reserve Bank of Minneapolis.

Stock, J. H. and M. W. Watson (2002): "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167–1179.

——— (2006): "Forecasting with Many Predictors," in *Handbook of Forecasting*, North Holland.

——— (2008a): "The Evolution of National and Regional Factors in U.S. Housing Construction," Princeton University.

——— (2008b): "Forecasting in Dynamic Factor Models Subjet to Structural Instability," Princeton University.

Sampling $\{F_t\}$

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 $\Psi_G(L)G_t = \Psi_G(L)\Lambda_F(L)F_t + \varepsilon_{Gt}$ or

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

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

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

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

or

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

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

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

$$\begin{aligned}
\vec{F}_{t|T}|\vec{F}^*_{t+1},\{\widetilde{G}_T\},\Xi_F &\sim N(\vec{F}_{t|t,\vec{F}^*_{t+1}},\vec{P}_{t|t,\vec{F}^*_{t+1}}) \\
\text{where } \vec{F}_{t|t,\,\vec{F}^*_{t+1}} &= \vec{F}_{t|t}+\vec{P}_{t|t}\vec{\Psi}^{*\prime}_F\,(\vec{\Psi}^*_F\,\vec{P}_{t|t}\,\vec{\Psi}^{*\prime}_F+\Sigma_F)^{-1}\,(\vec{F}^*_{t+1}-\vec{\Psi}^*_F\,\vec{F}_{t|t}) \\
\text{and } \vec{P}_{t|t,\vec{F}^*_{t+1}} &= \vec{P}_{t|t}-\vec{P}_{t|t}\,\vec{\Psi}^{*\prime}_F\,(\vec{\Psi}^*_F\,\vec{P}_{t|t}\,\vec{\Psi}^{*\prime}_F+\Sigma_F)^{-1}\,\vec{\Psi}^*_F\,\vec{P}_{t|t}.
\end{aligned} \tag{21}$$

where $\vec{F}^*_t$ and $\vec{\Psi}^*_F$ are the first K_F rows of $\vec{F}_t$ and $\vec{\Psi}_F$, respectively. Note also that we initialize the Kalman filter with the unconditional mean and variance of the states $\vec{F}$, i.e. $\vec{F}_{1|0}=E[\vec{F}]=0$ and $vec(\vec{P}_{1|0})=\left[I_{s_{F^*}}-\left(\vec{\Psi}_F\otimes\vec{\Psi}_F\right)\right]^{-1}vec(\vec{\Sigma}_F)$.

Sampling $\{G_t\}$

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 $\widetilde{X}_{bt}=\widetilde{\Lambda}_{Gb}(L)G_{bt}+\varepsilon_{Xbt},\forall\ b=1,\ldots,B$, where $\widetilde{X}_{bt}=\Psi_{Xb}(L)X_{bt}$ and $\widetilde{\Lambda}_{Gb}(L)=\Psi_{Xb}(L)\Lambda_{Gb}(L)$ is a $N_b\times K_b$ matrix polynomial of order $s^*_G=q_X+s_G$. Furthermore, $G_{bt}=\alpha_{Fbt}+\Psi_{G.b1}G_{bt-1}+\ldots\Psi_{G.bq_{Gb}}G_{bt-q_{Gb}}+\epsilon_{Gbt}$ where $\alpha_{Fbt}=\Psi_{Gb}(L)\Lambda_F(L)F_t,\qquad\forall\ b=1,\ldots,B$. Together, these two equations imply the following state-space form

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

$$\begin{pmatrix}G_{bt}\\ G_{bt-1}\\ \vdots\\ G_{bt-s^*_G}\end{pmatrix} = \begin{pmatrix}\alpha_{Fbt}\\ 0\\ \vdots\\ 0\end{pmatrix}+\begin{bmatrix}\Psi_{Gb1} & \cdots & \Psi_{Gbq_G} & 0 & \cdots & 0\\ I & 0 & \vdots & \vdots & & \vdots\\ \vdots & \ddots & \ddots & \vdots & & \vdots\\ 0 & \cdots & I & 0 & \cdots & 0\end{bmatrix}\begin{pmatrix}G_{bt-1}\\ G_{bt-2}\\ \vdots\\ G_{bt-s^*_G-1}\end{pmatrix}+\begin{pmatrix}\epsilon_{Gbt}\\ 0\\ \vdots\\ 0\end{pmatrix}$$

or

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

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

Denote Ξ_{Gb} the set of parameters $\{\vec{\Lambda}_{Gb},\vec{\Psi}_{Gb},\vec{\Sigma}_{Gb},\Sigma_{Xb}\}$. 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 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 $\vec{\alpha}_{Fbt}$, this implies the following steps:

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

We again initialize the filter with the unconditional mean and variance of the states $\vec{G}_b$, i.e. $\vec{G}_{b1|0} = E[\vec{G}_b]$ and $vec(\vec{P}_{1|0}) = vec(var(\vec{G}_b))$. Precisely, these are given by

$$
\begin{aligned}
E[\vec{G}_{bt}] &= E\left[\vec{\alpha}_{Fbt} + \vec{\Psi}_{Gb}\vec{G}_{bt-1} + \vec{\epsilon}_{Gbt}\right] = 0 \\
Var(\vec{G}_{bt}) &= Var(\vec{\alpha}_{Fbt} + \vec{\Psi}_{Gb}\vec{G}_{bt-1} + \vec{\epsilon}_{Gbt}) \\
&= Var(\vec{\alpha}_{Fbt}) + \vec{\Psi}_{Gb}Var(\vec{G}_{bt-1})\vec{\Psi}'_{Gb} + \vec{\Sigma}_{Gb} + 2\vec{\Psi}_{Gb}Cov(\vec{\alpha}_{Fbt}, \vec{G}_{bt-1})
\end{aligned}
$$

Altogether, we therefore have

$$
vec(Var(\vec{G}_b)) = \left[I - \left(\vec{\Psi}_{Gb} \otimes \vec{\Psi}_{Gb}\right)\right]^{-1}\left(\vec{\Sigma}_{\alpha_F} + \vec{\Sigma}_{Gb} + 2\vec{\Psi}_{Gb}\Sigma_{\alpha_F G_b}\right)
$$

where

$$
\begin{aligned}
\vec{\Sigma}_{\alpha_F} &= Var(\vec{\alpha}_{Fbt}) = \begin{bmatrix} \Sigma_{\alpha_F} & 0 \\ 0 & 0 \end{bmatrix} \\
\text{and} \qquad \Sigma_{\alpha_F} &= Var(\alpha_{Fbt}) = Var(\vec{\tilde{\Lambda}}_{Fb0}\vec{F}_t) = \vec{\tilde{\Lambda}}_{Fb0}Var(\vec{F}_t)\vec{\tilde{\Lambda}}'_{Fb0}
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\vec{\Sigma}_{\alpha_F G_b} &= Cov(\vec{\alpha}_{Fbt}, \vec{G}_{bt-1}) = \begin{bmatrix} \Sigma_{ab} & 0 \\ 0 & 0 \end{bmatrix}, \\
\text{with} \qquad \Sigma_{\alpha b} &= \vec{\tilde{\Lambda}}_{Fb1}Var(\vec{F}_{t-1})\vec{\Lambda}'_{Fb0}
\end{aligned}
$$

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:

$$
\vec{G}_{bt}|\widetilde{X}_{bt}, \vec{G}^*_{bt+1}, \Xi_{Gb} \sim N(\vec{G}_{bt|t,\vec{G}^*_{bt+1}}, \vec{P}_{Gbt|t,\vec{G}^*_{bt+1}}) \tag{22}
$$

where

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

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

Table A1: Data

	Block	Variables Ordered 1 and 2
1	CU	Capacity Utilization: Machinery (SA, Percent of Capacity)
		Capacity Utilization: Motor Vehicles and Parts (SA, Percent of Capacity)
2	IP	IP: Durable Consumer Goods (SA, 2002=100)
		IP: Nondurable Consumer Goods (SA, 2002=100)
3	ES	All Employees: Wholesale Trade (SA, Thous)
		Avg Wkly Earnings: Construction (SA, $/wk)
4	HS	Civilian Labor Force: Men: 25-54 Years (SA, Thous)
		Unemployment Rate: Full-Time Workers: Men (SA, %)
5	MS	ISM Mfg: PMI Composite Index (SA, 50+ = Econ Expand)
		Phila FRB Bus Outlook: General Activity, Current, Diffusion Index (SA,%)
6	DG	Mfrs' Inventories: Machinery (EOP, SA, Mil.$)
		Mfrs' Unfilled Orders: Machinery (EOP, SA, Mil.$)
7	RS	Retail Sales: General Merchandise Stores (SA, Mil.$)
		Retail Sales: Motor Vehicle Dealers (SA, Mil$)
8	WT	Merchant Wholesalers: Sales: Automotive (SA, Mil.$)
		Merchant Wholesalers: Sales: Apparel (SA, Mil.$)
9	AUTO	Domestic Car Retail Sales (SAAR, Mil. Units)
		Domestic Light Truck Retail Sales (SAAR, Mil. Units)
10	H- NE	Housing Starts: 1-Unit: Northeast (SAAR, Thous.Units)
		Housing Completions: 1-Unit: Northeast (SAAR, Thous.Units)
11	H- WEST	Housing Starts: 1-Unit: West (SAAR, Thous.Units)
		Housing Completions: 1-Unit: West (SAAR, Thous.Units)
12	H- CTL	Housing Starts: 1-Unit: Midwest (SAAR, Thous.Units)
		Housing Starts: 1-Unit: South (SAAR, Thous.Units)

Summary Statistics

b	T	N	IC_2	$R^2_{\widetilde{F}_1}$	$R^2_{\widetilde{F}_2}$	$AR_{\widetilde{F}_1}$	$AR_{\widetilde{F}_2}$
CU	191	25	1	0.210	0.092	0.112	0.031
IP	191	38	1	0.208	0.086	0.123	0.033
ES	191	72	2	0.189	0.132	-0.190	0.675
HS	191	85	8	0.113	0.081	-0.122	-0.446
MS	191	35	4	0.143	0.108	-0.062	-0.104
DG	191	60	2	0.115	0.088	0.273	0.302
RS	191	30	1	0.187	0.086	-0.333	-0.301
WT	191	53	1	0.093	0.064	-0.312	-0.115
AS	191	4	4	0.483	0.242	-0.360	-0.365
H- NE	191	8	8	0.190	0.181	-0.352	-0.531
H- WEST	191	7	7	0.240	0.196	-0.310	-0.199
H- CTL	191	18	0	0.119	0.091	-0.112	-0.246

Note: IC_2 is the Bai-Ng (2002) criteria for determining the number of factors. R^2_j is the j-th eigenvalue of $x'x$ divided by the sum of the eigenvalues. $AR_{\widetilde{F}_j}$ is the first order autocorrelation of j-th principal component of F.

Table A2: 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	$R^2_{G_{b,1}}$	$R^2_{Gb.2}$	$\widetilde{\Psi}_{\widetilde{G}.b1}$	$\widetilde{\Psi}_{\widetilde{G}.b2}$
CU	191	25	1	0.210	0.092	0.112	0.031
IP	191	38	1	0.208	0.086	0.123	0.033
ES	191	72	2	0.189	0.132	-0.190	0.675
HS	191	85	8	0.113	0.081	-0.122	-0.446
MS	191	35	4	0.143	0.108	-0.062	-0.104
DG	191	60	2	0.115	0.088	0.273	0.302
RS	191	30	1	0.187	0.086	-0.333	-0.301
WT	191	53	1	0.093	0.064	-0.312	-0.115
AS	191	4	4	0.483	0.242	-0.360	-0.365
H- NE	191	8	8	0.190	0.181	-0.352	-0.531
H- WEST	191	7	7	0.240	0.196	-0.310	-0.199
H- CTL	191	18	0	0.119	0.091	-0.112	-0.246

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

Table 1: A Three Level Housing Model

Posterior Means and Standard Deviations: $\widehat{\psi}_G$ and $\widehat{\psi}_F$

		Estimates		Standard Errors	
b	j	$\psi_{G.bj}$	$\widehat{\sigma}^2_{G.bj}$	$\psi_{G.bj}$	$\widehat{\sigma}^2_{G.bj}$
1	1	-0.007	0.366	0.177	0.098
1	2	-0.105	0.082	0.114	0.031
2	1	0.009	0.310	0.126	0.085
2	2	-0.251	0.091	0.128	0.038
3	1	0.110	0.116	0.179	0.066
3	2	-0.114	0.040	0.114	0.015
		ψ_F	σ^2_F		
	1	0.945	0.016		

Decomposition of Variance

	Estimates				Standard Errors			
block	σ^2_X	share$_F$	share$_G$	share$_X$	σ^2_X	share$_F$	share$_G$	share$_X$
	Data							
1	1.490	0.173	0.237	0.590	2.299	0.133	0.059	0.087
2	1.742	0.106	0.230	0.664	5.924	0.103	0.051	0.064
3	1.368	0.140	0.118	0.742	1.269	0.101	0.032	0.077

Table 2: A Six Block Three Level Model for Production:

$$\begin{aligned} G_{bt} &= \Lambda_{F.b}(L)F_t + e_{Gbt} \\ \Psi_{F.r}(L)F_{rt} &= \epsilon_{Frt}, \quad r = 1, \dots K_F \\ \Psi_{G.bj}(L)e_{Gbjt} &= \epsilon_{Gbjt}, \quad j = 1, \dots K_{Gb}. \end{aligned}$$

Block	j	$\widehat{\Psi}_{G.bj}$	$\widehat{\sigma}^2_{\epsilon bj}$	S.E	
CU: 1	1	0.272	0.031	0.114	0.026
CU: 1	2	0.288	0.072	0.166	0.033
IP: 2	1	0.322	0.043	0.117	0.020
IP: 2	2	0.098	0.064	0.107	0.018
ES: 3	1	0.013	0.027	0.185	0.010
ES: 3	2	-0.034	0.027	0.185	0.010
HS: 4	1	-0.067	0.103	0.100	0.021
HS: 4	2	-0.021	0.059	0.101	0.013
MS: 5	1	0.381	0.860	0.130	0.103
MS: 5	2	0.087	0.106	0.099	0.023
DG: 6	1	-0.101	0.039	0.173	0.009
DG: 6	2	0.041	0.041	0.175	0.009
Factor		$\Psi_{F.k}$	$\widehat{\sigma}^2_{F.k}$	S.E.	
1		0.795	0.020	0.068	0.008

Principal Component Estimates

	N	$R^2_{F.1}$	$R^2_{F.2}$	$\widetilde{\Psi}_{F.1}$	$\widetilde{\Psi}_{F.2}$
$\widetilde{F}_t(\widetilde{G}_t)$	14	.205	.133	.074	.272
$\widetilde{F}_t$	315	.210	.208	.187	.339

Decomposition of Variance

	Estimates				Standard Errors			
block	σ^2_X	share$_F$	share$_G$	share$_X$	σ^2_X	share$_F$	share$_G$	share$_X$
1 CU:	1.235	0.197	0.144	0.659	0.120	0.048	0.021	0.042
2 IP:	1.318	0.245	0.151	0.604	0.153	0.054	0.019	0.043
3 ES:	1.116	0.201	0.138	0.661	0.112	0.053	0.020	0.037
4 HS:	1.061	0.021	0.170	0.809	0.029	0.015	0.011	0.013
5 MS:	1.130	0.044	0.244	0.712	0.053	0.025	0.020	0.018
6 DG:	1.033	0.062	0.141	0.798	0.040	0.028	0.013	0.024

Table 3: Correlation Between $\widehat{G}_{bkt}$ and $\widetilde{e}_{rt|\widehat{F}}$

r	b	k	ρ	r	b	k	ρ
1	3	2	0.235	4	1	2	0.109
2	3	2	0.716	4	6	1	0.233
2	5	1	0.124	5	4	2	0.139
2	6	2	0.196	6	2	2	0.138
3	4	1	0.616	7	5	2	0.224

Table 4: A Nine Block Four Level Model for Real Activity

b	j	$\widehat{\Psi}_{G.bj}$	$\widehat{\sigma}^2_{\epsilon bj}$	S.E	
1	1	0.302	0.101	0.140	0.024
2	1	0.351	0.037	0.193	0.015
2	2	0.342	0.019	0.328	0.007
3	1	0.182	0.063	0.173	0.024
Factor		$\Psi_{F.k}$	$\widehat{\sigma}^2_{F.k}$	S.E.	
1		0.895	0.013	0.049	0.007

Decomposition of Variance

block	sub-block	σ^2_X	share$_F$	share$_G$	share$_H$	share$_X$
Estimates						
1	CU	1.239	0.104	0.137	0.169	0.590
1	IP	1.157	0.099	0.129	0.153	0.619
1	MS	1.050	0.027	0.033	0.237	0.703
1	DG	0.975	0.022	0.027	0.172	0.779
2	ES	1.032	0.051	0.103	0.205	0.641
2	HS	1.026	0.018	0.037	0.177	0.767
3	RS	1.091	0.045	0.072	0.213	0.669
3	WT	1.007	0.015	0.023	0.164	0.798
3	AU	1.050	0.040	0.066	0.552	0.342
Standard Errors						
1	CU	0.525	0.063	0.033	0.027	0.040
1	IP	0.404	0.061	0.029	0.020	0.041
1	MS	0.076	0.027	0.012	0.018	0.023
1	DG	0.072	0.023	0.008	0.013	0.021
2	ES	0.167	0.044	0.026	0.027	0.033
2	HS	0.058	0.023	0.013	0.014	0.022
3	RS	0.509	0.044	0.020	0.022	0.032
3	WT	0.147	0.022	0.010	0.015	0.022
3	AU	0.332	0.046	0.034	0.051	0.027

Figure 1: “True” and Estimated Factors for Housing

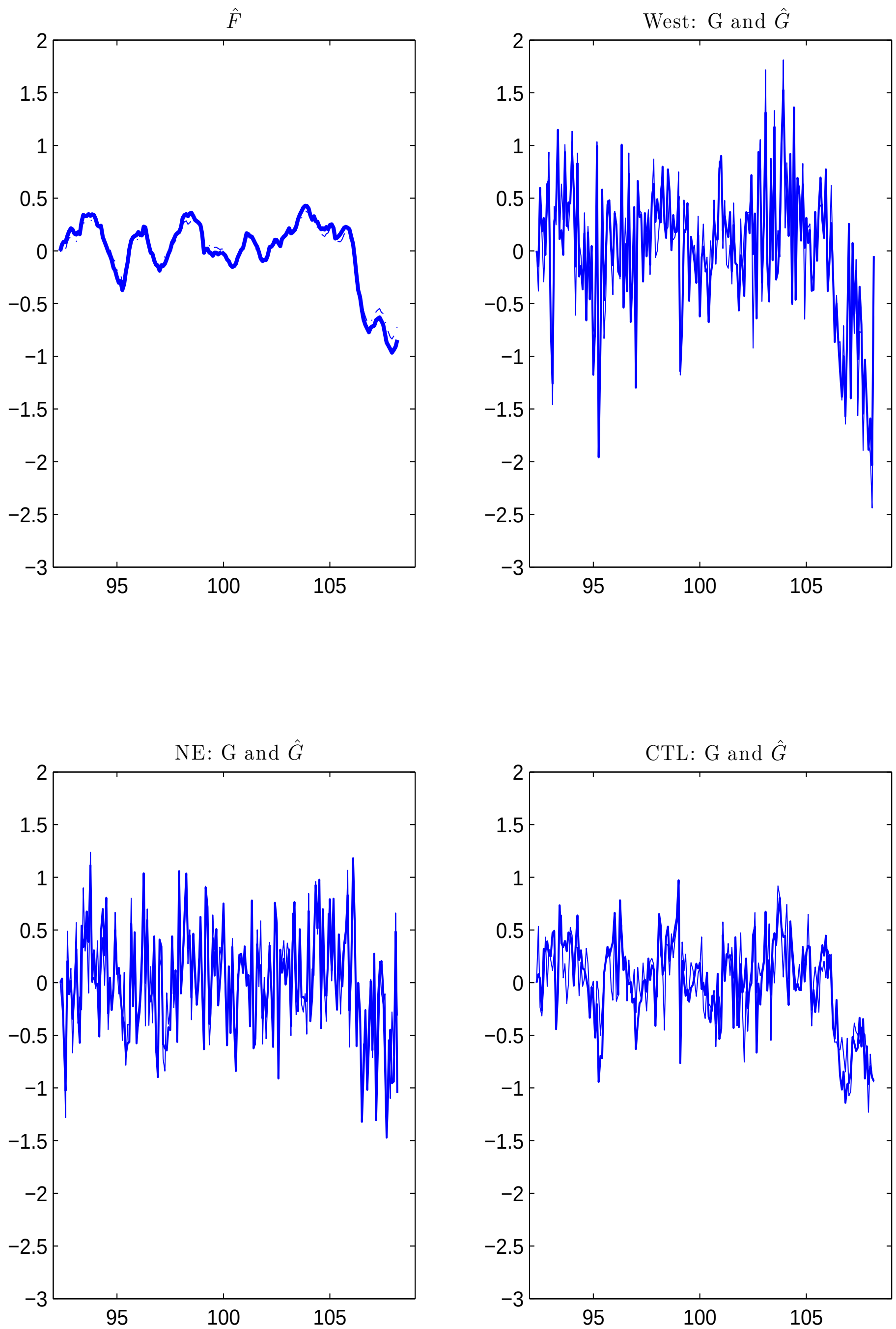

Figure 2: 6-Block, 3 Level Model of Output

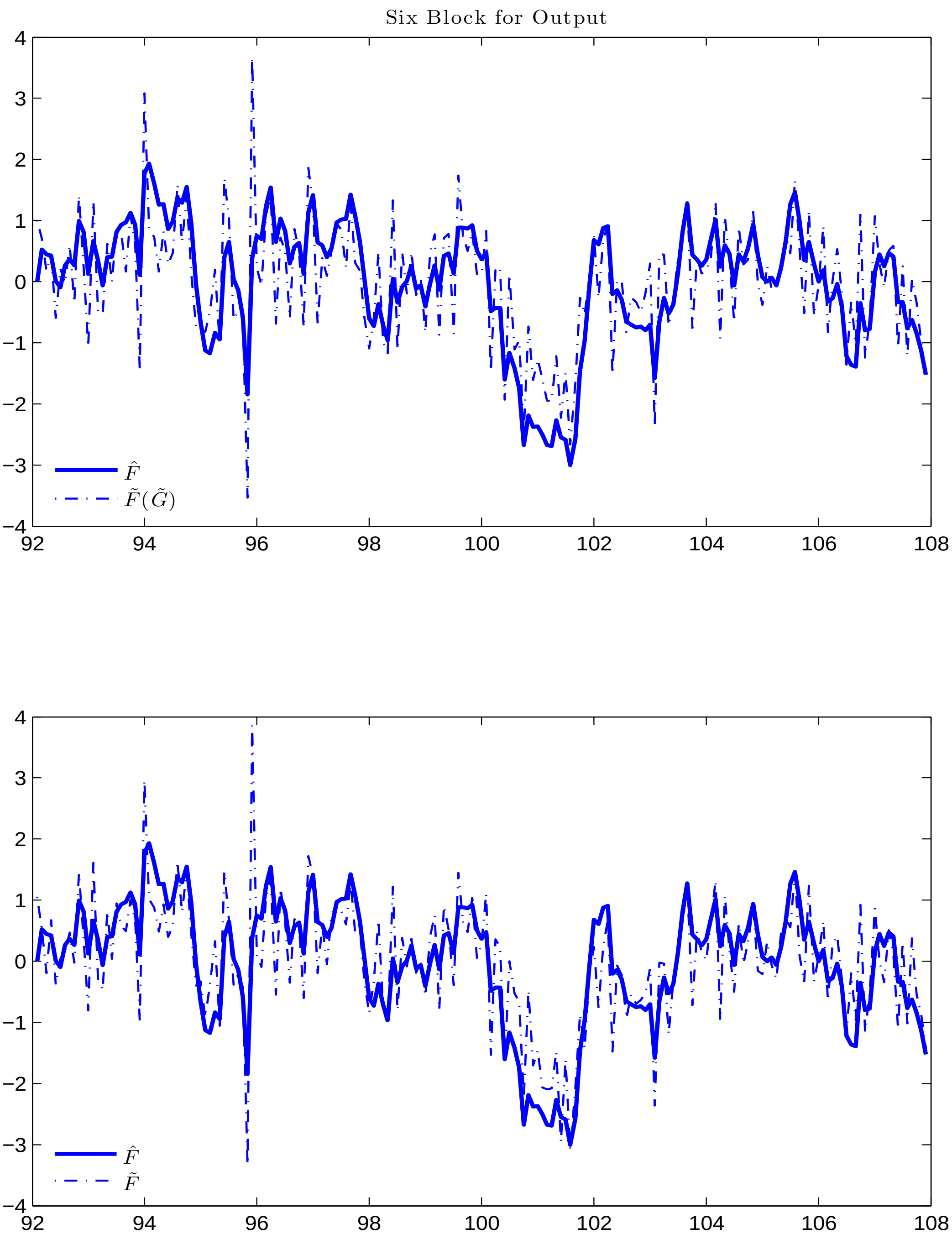

Figure 3: 9-Block, 4 Level Model of Real Economic Activity