Dynamic Hierarchical Factor Models

Emanuel Moench$^2$     Serena Ng$^1$     Simon Potter$^2$

$^1$Columbia University  $^2$New York Fed

December 2009
Motivation

A multi-level (hierarchical) factor model:

- A large panel of data organized by $B$ blocks,
  - e.g. Production, Employment, Demand, Housing, ...
- each block $b$ has $N_b$ series, $N_b$ large
- $N = \sum_{b=1}^{B} N_b$
Motivation

A multi-level (hierarchical) factor model:

- A large panel of data organized by $B$ blocks,
  - e.g. Production, Employment, Demand, Housing, ...
  - each block $b$ has $N_b$ series, $N_b$ large
- $N = \sum_{b=1}^{B} N_b$
- Each block can be divided into sub-blocks
  - e.g. sub-blocks of Demand: Retail Sales, Auto Sales, Wholesale Trade
Motivation

A multi-level (hierarchical) factor model:

- A large panel of data organized by $B$ blocks,
  - e.g. Production, Employment, Demand, Housing, …
  - each block $b$ has $N_b$ series, $N_b$ large
- $N = \sum_{b=1}^{B} N_b$
- Each block can be divided into sub-blocks
  - e.g. sub-blocks of Demand: Retail Sales, Auto Sales, Wholesale Trade
- Within block variations due to block-level factors
- Between block variations due to common factors
- Idiosyncratic noise
A Three Level Model

Level 1: For $b = 1, \ldots, B$, $i = 1, \ldots, N_b$,

$$X_{bit} = \Lambda_{G.bi}(L)G_{bt} + e_{X_{bit}},$$

$$\psi_{X.bi}(L)e_{X_{bit}} = \epsilon_{X_{bit}},$$
A Three Level Model

**Level 1:** For $b = 1, \ldots, B$, $i = 1, \ldots, N_b$,

\[
X_{bit} = \Lambda_{G.bi}(L) G_{bt} + e_{Xbit},
\]

\[
\psi_{X.bi}(L) e_{Xbit} = \epsilon_{Xbit},
\]

**Level 2:** For $j = 1, \ldots k_{Gb}$

\[
G_{bjt} = \Lambda_{F.bj}(L) F_t + e_{Gbjt},
\]

\[
\psi_{G.bj}(L) e_{Gbjt} = \epsilon_{Gbjt}.
\]
A Three Level Model

Level 1: For \( b = 1, \ldots, B \), \( i = 1, \ldots, N_b \),

\[
X_{bit} = \Lambda_{G.bi}(L)G_{bt} + e_{Xbit}, \\
\psi_{X.bi}(L)e_{Xbit} = \epsilon_{Xbit},
\]

Level 2: For \( j = 1, \ldots k_{Gb} \)

\[
G_{bjt} = \Lambda_{F.bj}(L)F_t + e_{Gbjt}, \\
\psi_{G.bj}(L)e_{Gbjt} = \epsilon_{Gbjt}.
\]

Level 3: For \( r = 1, \ldots k_F \)

\[
\psi_{F.r}(L)F_{rt} = \epsilon_{Fr},
\]

\( \epsilon_{Xbit}, \epsilon_{Gbjt}, \epsilon_{Fkt} \sim iid(0, \sigma^2_{Xbi}, \sigma^2_{Gbj}, \sigma^2_{Fr}). \)
A Four Level Model

For $s = 1, \ldots, S_b$, $b = 1, \ldots, B$, $i = 1, \ldots, N_b$:

\[
Z_{bsit} = \Lambda_{H.bsi}(L)H_{bst} + e_{Zbsit}, \quad \text{Individual Series}
\]
\[
H_{bst} = \Lambda_{G.bs}(L)G_{bt} + e_{Hbst}, \quad \text{Subblock Factors}
\]
\[
G_{bt} = \Lambda_{F.b}(L)F_{t} + e_{Gbt}, \quad \text{Block factors}
\]
\[
\psi_{F.r}(L)F_{rt} = \epsilon_{Frt}, \quad \text{Aggregate factors}
\]

with dynamics

\[
\psi_{Z.bsi}(L)e_{Zbsit} = \epsilon_{Zbsit}
\]
\[
\psi_{H.bs}(L)e_{Hbst} = \epsilon_{Hbst}
\]
\[
\psi_{G.b}(L)e_{Gbt} = \epsilon_{Gbt}
\]
Why another factor model?

1) Block structure arises naturally in many economic and financial analyses:

- real activity: production, employment, demand, housing
- Global, regional, and country level variations
- Country, regions, state level variations
- Aggregate, industry, firm level variations in stock returns
2) Can put structure on the model through

- Ordering of the variables within block
- Ordering of the blocks

\[ X_{bt} = \Lambda_{Gb.0} G_{bt} + \ldots + \Lambda_{Gb.sGb} G_{b,t-sGb} + eX_{bt} \]

where

\[ \Lambda_{G.b0} = \begin{bmatrix} 1 & 0 & 0 \\ x & \ddots & 0 \\ x & x & 1 \\ \lambda_{G.b01} & \cdots & \lambda_{G.b0k_b} \end{bmatrix} \]

Easy interpretation of factors
3) Addresses a limitation of level two models:

\[ k_G = k_F = 1: \]

\[
x_{ibt} = \lambda_{G.ib}(\lambda_{F.b}F_t + e_{G.bt}) + e_{X.ibt}
\]

\[
= \lambda_{ib}F_t + v_{ibt}
\]

\[
v_{ibt} = \lambda_{G.ib}e_{G.bt} + e_{X.ibt}.
\]

Ignoring block level variations gives a level 2 model:

\[
x_{it} = \lambda_iF_t + v_{it}
\]

with \( E(v_{it}v_{jt}) \neq 0 \) if \( i,j \) both belong to \( b \).

A multi-level model controls for these ‘quasi-common’ variations.
4) State Space Framework

- Data sampled at mixed frequencies
- Missing values
- Internally coherent (no auxiliary forecasting model necessary)
5) Advantages and Uses

- Allows block and aggregate level analysis but still achieves dimension reduction
- Jointly estimates block level and aggregate factors
- Can be used for monitoring, counterfactuals, assess relative importance of shocks etc.
Application: Monitoring Real Activity in the US

- Data are released in various blocks throughout the month
- Releases broadly correspond to economic categories
- Block level factors are of independent interest
- Real time monitoring of $F_t$ and $G_{bt}$
- More manageable than monitoring hundreds of series
Motivation
Why another factor model
Related Literature
Level 3
Results
Level 4

Related Work

1. Diebold, Li, Yue (JOE 2008): Three level model for international bond yields

- Two step: estimate country level factors using OLS (with loadings fixed), then estimate global factors via MCMC.
- Limitations:
  - single factor at block and global levels
  - Information from global factors not taken into account when sampling country (block-level) factors.
  - Global factors do not account for sampling uncertainty of block factors.

Our one-step estimation solves both issues.
2 Kose-Otrok-Whiteman (AER 2003): three level model

- For each unit \( i \) in country \( b \):
  \[
  x_{bit} = c_i F_t + d_{bi} e_{Gbt} + e_{bit}
  \]

- Top down vs. bottom up (\( G_b \) vs \( e_{Gb} \))
- Single factors
- Static loadings
- \((N \cdot k_F + N \cdot k_G)\) vs \((k_G \cdot k_F + N \cdot k_G)\) parameters.
- Hierarchical structure: \( k_G \ll N \)
- For each \( i \), need to invert a \( T \times T \) matrix at each draw.
3 Hierarchical loadings vs. hierarchical factors

- Common factors across blocks, loadings differ by blocks
  - Spatial factor models: loadings vary by distance
- Not a natural way of analyzing macroeconomic data
- All shocks are global, sensitivity to global shocks differ by regions
4 Principal components

i One step estimation of $F_t$ (level two)

ii One step estimation of $F_t$ and $e_{Gbt}$ (but no $G_{bt}$)

iii Sequential estimation: $\tilde{F}_t(\tilde{G}_t)$:
  
  - For each $b$, first estimate $G_{bt}$, then estimate $F_t$ from $\tilde{G}_t$
  - Ignore dynamic dependence of $G_{bt}$ on $F_t$.
  - Needs auxiliary equations for forecasting

iv Block level dynamic principal components?

Require $N$ and $T$ large.
Many other seemingly related state space models.....

Unique features of our dynamic hierarchical model:

- Coherent treatment of factors at different levels
- Produce factor estimates at both the block-level and aggregate levels
- Multiple factors at each level
- Hierarchical structure of factors, not loadings
- Analyze up to 4 levels, each with possibly multiple factors
- Can handle (but does not require) large $N$ or $T$. 
The State Space Form

Block-Level Factors:

\[ G_{bt} = \Lambda_{F.b0} F_t + \Lambda_{F.b1} F_{t-1} + \ldots + \Lambda_{F.bl_F} F_{t-l_F} + e_{Gbt}, \]
\[ \Psi_{G.b}(L)e_{Gbt} = \epsilon_{Gbt}. \]

implies (pseudo) measurement equation

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

Transition equation:

\[ \Psi_F(L)F_t = \epsilon_{Ft}, \]
Observed data:

\[ X_{bt} = \Lambda_{Gb.0} G_{bt} + \ldots + \Lambda_{Gb.sGb} G_{b,t-sGb} + \epsilon_{Xbt} \]

\[ \psi_{X.bi}(L) e_{Xbit} = \epsilon_{Xbit}, \]

implies the individual level measurement equation

\[ \psi_{X.b}(L) X_{bt} = \psi_{X.b}(L) \Lambda_{G.b}(L) G_{bt} + \epsilon_{Xbt}. \]

Measurement equation from level 2 becomes the transition equation for state variable in level 1

\[ G_{bt} = \alpha_{F.bt} + \psi_{G.b1} G_{bt-1} + \ldots + \psi_{G.bqGb} G_{bt-qGb} + \epsilon_{Gbt}. \]
Observed data:

\[ X_{bt} = \Lambda G_{b0}G_{bt} + \ldots + \Lambda G_{sbGb}G_{bt-sGb} + e_{Xbt} \]

\[ \psi_{X.bi}(L)e_{Xbit} = \epsilon_{Xbit}, \]

implies the individual level measurement equation

\[ \psi_{X.b}(L)X_{bt} = \psi_{X.b}(L)\Lambda G.b(L)G_{bt} + \epsilon_{Xbt}. \]

Measurement equation from level 2 becomes the transition equation for state variable in level 1

\[ G_{bt} = \alpha_{F.bt} + \psi_{G.b1}G_{bt-1} + \ldots + \psi_{G.bqGb}G_{bt-qGb} + \epsilon_{Gbt}. \]

with time-varying intercept

\[ \alpha_{F.bt} = \psi_{G.b}(L)\Lambda F.b(L)F_t. \]
Identification

- \( \text{var}(\varepsilon_{Gb}) \) and \( \text{var}(\varepsilon_F) \) are diagonal
- Block factor loadings

\[
\Lambda_{G0} = \begin{bmatrix}
1 & 0 & 0 \\
x & \ddots & 0 \\
x & x & 1 \\
\lambda_{Gb01} & \cdots & \lambda_{Gb0k_b}
\end{bmatrix}
\]

- Common factor loadings

\[
\Lambda_{F0} = \begin{bmatrix}
1 & 0 & 0 \\
x & \ddots & 0 \\
x & x & 1 \\
\lambda_{F.01} & \cdots & \lambda_{F.0K_F}
\end{bmatrix}
\]
MCMC

Let $\Sigma = (\Sigma_F, \Sigma_G, \Sigma_X)$, $\Psi = (\Psi_F, \Psi_G, \Psi_X)$, $\Lambda = (\Lambda_G, \Lambda_F)$

1. Use PCs as initial estimates of $\{G_t\}$ and $\{F_t\}$, get initial values for $\Lambda$, $\Psi$, $\Sigma$

2. Conditional on $\Lambda$, $\Psi$, $\Sigma$ and $\{F_t\}$: draw $\{G_{bt}\}$ block by block using Carter-Kohn algorithm modified to allow for time varying intercept

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

4. Conditional on $\{F_t\}$ and $\{G_t\}$: draw $\Lambda$, $\Psi$, and $\Sigma$ assuming conjugate priors
Step 2: allows lower level factors to depend on the factors at the next level.

Level two transition equation: \( \alpha_{F.bt} = \Psi_{Gb}(L) \Lambda_F(L) F_t \)

\[
G_{bt} = \alpha_{F.bt} + \Psi_{Gb.1} G_{bt-1} + \ldots \Psi_{Gb.q_{Gb}} G_{bt-q_{Gb}} + \epsilon_{Gb_t}
\]

- The time-varying intercept \( \alpha_{F.bt} \) is known given \( \Psi, \Lambda, \{F_t\} \):
- Modify standard updating and smoothing equations for \( G_{bt} \) to take this into account.
- Any filtering/sampling method for linear state space models can be adapted to hierarchical models this way.
Data Release Calendar June 2009

- June 1
  - ISM Manufacturing Survey (ISM)
  - Auto Sales (AUTO)

- June 2
  - Wholesale Trade (WT)

- June 3
  - Manufacturing Shipments, Inventories, and Orders (DG)

- June 6
  - Establishment Survey (ES)
  - Household Survey (HS)

- June 12
  - Retail Sales (RS)

- June 15
  - Housing Starts (H-Starts)
  - Industrial Production (IP)
  - Capacity Utilization (CU)

- June 17
  - Philadelphia Business Outlook Survey (PhilaFed)

- June 19
  - Existing Home Sales (H-ExistSales)

- June 25
  - New Home Sales (H-NewSales)

- June 26
  - Chicago Midwest Mfg. Survey (ChicagoFed)

Motivation Why another factor model

Related Literature

Level 3

Results

Level 4
A Three Level Model: 315 series

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

Parameters:

- $k_F = 1$, $k_G = 2$
- $s_F = 2$, $s_G = 2$
- $q_X = q_G = q_F = 1$
### Table 4: Level 3, 6 Block Model

<table>
<thead>
<tr>
<th>Block</th>
<th>j</th>
<th>$\hat{\Psi}_{G, bj}$</th>
<th>$\hat{\sigma}^2_{\epsilon, bj}$</th>
<th>S.E</th>
</tr>
</thead>
<tbody>
<tr>
<td>CU: 1</td>
<td>1</td>
<td>0.373</td>
<td>0.064</td>
<td>0.113</td>
</tr>
<tr>
<td>CU: 1</td>
<td>2</td>
<td>-0.122</td>
<td>0.057</td>
<td>0.091</td>
</tr>
<tr>
<td>IP: 2</td>
<td>1</td>
<td>0.170</td>
<td>0.015</td>
<td>0.110</td>
</tr>
<tr>
<td>IP: 2</td>
<td>2</td>
<td>-0.140</td>
<td>0.047</td>
<td>0.089</td>
</tr>
<tr>
<td>ES: 3</td>
<td>1</td>
<td>0.052</td>
<td>0.015</td>
<td>0.137</td>
</tr>
<tr>
<td>ES: 3</td>
<td>2</td>
<td>-0.160</td>
<td>0.031</td>
<td>0.115</td>
</tr>
<tr>
<td>HS: 4</td>
<td>1</td>
<td>0.198</td>
<td>0.137</td>
<td>0.095</td>
</tr>
<tr>
<td>HS: 4</td>
<td>2</td>
<td>-0.069</td>
<td>0.056</td>
<td>0.096</td>
</tr>
<tr>
<td>MS: 5</td>
<td>1</td>
<td>0.436</td>
<td>0.824</td>
<td>0.128</td>
</tr>
<tr>
<td>MS: 5</td>
<td>2</td>
<td>0.059</td>
<td>0.111</td>
<td>0.093</td>
</tr>
<tr>
<td>DG: 6</td>
<td>1</td>
<td>-0.013</td>
<td>0.030</td>
<td>0.172</td>
</tr>
<tr>
<td>DG: 6</td>
<td>2</td>
<td>-0.009</td>
<td>0.030</td>
<td>0.175</td>
</tr>
<tr>
<td>Factor</td>
<td></td>
<td>$\Psi_{F,k}$</td>
<td>$\hat{\sigma}^2_{F,k}$</td>
<td>S.E.</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.880</td>
<td>0.061</td>
<td>0.040</td>
</tr>
</tbody>
</table>
## Decomposition of Variance Estimates

<table>
<thead>
<tr>
<th>block</th>
<th>share$_F$</th>
<th>share$_G$</th>
<th>share$_X$</th>
<th>share$_F$</th>
<th>share$_G$</th>
<th>share$_X$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 CU:</td>
<td>0.303</td>
<td>0.144</td>
<td>0.553</td>
<td>0.069</td>
<td>0.021</td>
<td>0.055</td>
</tr>
<tr>
<td>2 IP:</td>
<td>0.321</td>
<td>0.131</td>
<td>0.549</td>
<td>0.075</td>
<td>0.021</td>
<td>0.058</td>
</tr>
<tr>
<td>3 ES:</td>
<td>0.279</td>
<td>0.114</td>
<td>0.607</td>
<td>0.073</td>
<td>0.020</td>
<td>0.056</td>
</tr>
<tr>
<td>4 HS:</td>
<td>0.081</td>
<td>0.150</td>
<td>0.769</td>
<td>0.034</td>
<td>0.013</td>
<td>0.026</td>
</tr>
<tr>
<td>5 MS:</td>
<td>0.117</td>
<td>0.222</td>
<td>0.661</td>
<td>0.056</td>
<td>0.033</td>
<td>0.031</td>
</tr>
<tr>
<td>6 DG:</td>
<td>0.101</td>
<td>0.123</td>
<td>0.777</td>
<td>0.044</td>
<td>0.014</td>
<td>0.035</td>
</tr>
</tbody>
</table>
**Figure:** 6 Block Model of Real Activity:
Is $\tilde{F}_t$ picking up block-level common variations?

- Bai and Ng (2002): $IC_2$ finds 2 factors common to $\tilde{G}_t$
- for $r = 1, \ldots, k_F$, regress $\tilde{F}_{rt}$ on $\hat{F}_t$. Let $\tilde{e}_{rt}$ be the residuals
- regress $\tilde{e}_{rt}$ on $\hat{G}_t$
- $R^2$ measures variations in $\tilde{F}_{kt}$ that are not genuinely common: orthogonal to our $\hat{F}_t$ but correlated with $G_{bt}$. 
Table 5: Correlation Between $\hat{G}_{bkt}$ and $\tilde{e}_{rt|\hat{F}}$

<table>
<thead>
<tr>
<th>k</th>
<th>b</th>
<th>j</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0.17</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0.20</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0.32</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>1</td>
<td>0.13</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>1</td>
<td>0.27</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td>0.16</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>0.60</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2</td>
<td>0.31</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>2</td>
<td>0.15</td>
</tr>
</tbody>
</table>
## Four Level Model: 447 Series

<table>
<thead>
<tr>
<th>Block</th>
<th>subblock</th>
<th>$N$</th>
<th>$K_{Gb}$</th>
<th>$K_{Hbs}$</th>
<th>Variable Ordered First</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>CU</td>
<td>25</td>
<td>1</td>
<td>2</td>
<td>Capacity Utilization</td>
</tr>
<tr>
<td></td>
<td>IP</td>
<td>38</td>
<td>1</td>
<td>2</td>
<td>IP: Durables</td>
</tr>
<tr>
<td></td>
<td>DG</td>
<td>60</td>
<td>1</td>
<td>2</td>
<td>Manufacturers’ Inventories: Machinery</td>
</tr>
<tr>
<td>Employment</td>
<td>ES</td>
<td>82</td>
<td>1</td>
<td>2</td>
<td>All Employees: Wholesale</td>
</tr>
<tr>
<td></td>
<td>HS</td>
<td>92</td>
<td>1</td>
<td>2</td>
<td>Civilian Labor Force: Men</td>
</tr>
<tr>
<td>Demand</td>
<td>RS</td>
<td>30</td>
<td>1</td>
<td>2</td>
<td>Retail Sales: General Merchandise</td>
</tr>
<tr>
<td></td>
<td>WS</td>
<td>54</td>
<td>1</td>
<td>2</td>
<td>Merchant Wholesalers: Sales: Automotive</td>
</tr>
<tr>
<td></td>
<td>AUTO</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>Domestic Car Retail Sales</td>
</tr>
<tr>
<td>Housing</td>
<td>H-STARTS</td>
<td>24</td>
<td>1</td>
<td>2</td>
<td>Housing Starts: 1-Unit: West</td>
</tr>
<tr>
<td></td>
<td>H-NEWSALES</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>New 1-Family Houses Sold</td>
</tr>
<tr>
<td></td>
<td>H-EXISTSALES</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>NAR Total Existing Home</td>
</tr>
<tr>
<td>Mfg Surveys</td>
<td>ISM</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>ISM Mfg: PMI Composite</td>
</tr>
<tr>
<td></td>
<td>PHILAFED</td>
<td>21</td>
<td>1</td>
<td>1</td>
<td>Phila FRB Bus Outlook: General</td>
</tr>
<tr>
<td></td>
<td>CHICFED</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>Chicago FRB: Midwest Mfg</td>
</tr>
</tbody>
</table>
Four Level Model for Real Activity with 5 Blocks and 14 Subblocks: $\hat{F}$ and all $\hat{G}$
## Decomposition of Variance

<table>
<thead>
<tr>
<th>block</th>
<th>sub-block</th>
<th>share_{F}</th>
<th>share_{G}</th>
<th>share_{H}</th>
<th>share_{X}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>IP</td>
<td>0.090</td>
<td>0.115</td>
<td>0.176</td>
<td>0.619</td>
</tr>
<tr>
<td>Output</td>
<td>CU</td>
<td>0.089</td>
<td>0.113</td>
<td>0.161</td>
<td>0.637</td>
</tr>
<tr>
<td>Output</td>
<td>DG</td>
<td>0.022</td>
<td>0.026</td>
<td>0.180</td>
<td>0.773</td>
</tr>
<tr>
<td>Employment</td>
<td>HS</td>
<td>0.042</td>
<td>0.090</td>
<td>0.226</td>
<td>0.642</td>
</tr>
<tr>
<td>Employment</td>
<td>ES</td>
<td>0.015</td>
<td>0.033</td>
<td>0.171</td>
<td>0.782</td>
</tr>
<tr>
<td>Demand</td>
<td>RS</td>
<td>0.030</td>
<td>0.073</td>
<td>0.212</td>
<td>0.685</td>
</tr>
<tr>
<td>Demand</td>
<td>WT</td>
<td>0.012</td>
<td>0.028</td>
<td>0.153</td>
<td>0.806</td>
</tr>
<tr>
<td>Demand</td>
<td>AUTO</td>
<td>0.018</td>
<td>0.049</td>
<td>0.393</td>
<td>0.539</td>
</tr>
<tr>
<td>Housing</td>
<td>H-starts</td>
<td>0.005</td>
<td>0.069</td>
<td>0.176</td>
<td>0.750</td>
</tr>
<tr>
<td>Housing</td>
<td>H-newsales</td>
<td>0.002</td>
<td>0.030</td>
<td>0.285</td>
<td>0.683</td>
</tr>
<tr>
<td>Housing</td>
<td>H-existsales</td>
<td>0.003</td>
<td>0.038</td>
<td>0.626</td>
<td>0.333</td>
</tr>
<tr>
<td>Mfg Surveys</td>
<td>ISM</td>
<td>0.035</td>
<td>0.161</td>
<td>0.248</td>
<td>0.556</td>
</tr>
<tr>
<td>Mfg Surveys</td>
<td>PhilaFed</td>
<td>0.010</td>
<td>0.045</td>
<td>0.197</td>
<td>0.747</td>
</tr>
<tr>
<td>Mfg Surveys</td>
<td>ChicagoFed</td>
<td>0.022</td>
<td>0.071</td>
<td>0.461</td>
<td>0.446</td>
</tr>
</tbody>
</table>
Monitoring in a data rich environment

• Update state of the block as new information in sub-block arrives
• ‘Predict’ state of blocks for which new data are not yet released
• Use updated and predicted block factors to update aggregate factors
Figure: Four Level Model with 5 Blocks and 14 Subblocks
Figure: Demand Block

$\hat{G}_3$: Demand

- Monthly update
- Continuous update
Figure: Housing

$\hat{G}_4$: Housing

- Monthly update
- Continuous update
Figure: Manufacturing

$G_5$: Mfg Surveys

Monthly update
Continuous update
A Housing Model

Three features of the model:

i) Distinguishes regional from national variations
   • 4 Census regions: North-East, West, Mid-West, South.

ii) Distinguishes house price from housing market variations, i.e. use data on prices and volume.

iii) Combines data that are sampled at monthly and quarterly frequencies and available over different time spans.

iv) Allows observed factors.
### Regional Series

<table>
<thead>
<tr>
<th>Price data</th>
<th>Source</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Sales Price of Single Family Existing Homes</td>
<td>NAR</td>
<td>mly</td>
</tr>
<tr>
<td>Single Family Median Home Sales Price</td>
<td>CENSUS</td>
<td>qly</td>
</tr>
<tr>
<td>Average Existing Home Prices</td>
<td>NAR</td>
<td>qly</td>
</tr>
<tr>
<td>Average New Home Prices</td>
<td>NAR</td>
<td>qly</td>
</tr>
<tr>
<td>Conventional Mortgage Home Price Index</td>
<td>FHLMC</td>
<td>qly</td>
</tr>
<tr>
<td>OFHEO Purchase-only Index</td>
<td>OFHEO</td>
<td>mly</td>
</tr>
<tr>
<td>OFHEO Home Prices</td>
<td>OFHEO</td>
<td>qly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Volume data</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>New 1-Family Houses Sold</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>New 1-Family Houses For Sale</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>Single-Family Housing Units Under Construction</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>Multifamily Units Under Construction</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>Homeownership Rate</td>
<td>CENSUS</td>
<td>qly</td>
</tr>
<tr>
<td>Homeowner Vacancy Rate</td>
<td>CENSUS</td>
<td>qly</td>
</tr>
<tr>
<td>Rental Vacancy Rate</td>
<td>CENSUS</td>
<td>qly</td>
</tr>
</tbody>
</table>
### National Series

<table>
<thead>
<tr>
<th>Price data</th>
<th>Source</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Sales Price of Single Family Existing Homes</td>
<td>NAR</td>
<td>mly</td>
</tr>
<tr>
<td>Median Sales Price of Single Family New Homes</td>
<td>Census</td>
<td>mly</td>
</tr>
<tr>
<td>Single Family Median Home Sales Price</td>
<td>CENSUS</td>
<td>qly</td>
</tr>
<tr>
<td>Average Existing Home Prices</td>
<td>NAR</td>
<td>mly</td>
</tr>
<tr>
<td>Average New Home Prices</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>S&amp;P/Case-Shiller Home Price Index</td>
<td>S&amp;P</td>
<td>qly</td>
</tr>
<tr>
<td>Conventional Mortgage Home Price Index</td>
<td>FHLMC</td>
<td>qly</td>
</tr>
<tr>
<td>OFHEO Purchase-only Index</td>
<td>OFHEO</td>
<td>mly</td>
</tr>
<tr>
<td>OFHEO Home Prices</td>
<td>OFHEO</td>
<td>qly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Volume data</th>
<th>Source</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>New 1-Family Houses For Sale</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>Housing Units Authorized by Permit: 1-Unit</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>Multifamily Units Under Construction</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>Multifamily Permits US</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>Multifamily Starts US</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>Multifamily Completions</td>
<td>CENSUS</td>
<td>mly</td>
</tr>
<tr>
<td>Homeowner Vacancy Rate</td>
<td>CENSUS</td>
<td>qly</td>
</tr>
<tr>
<td>Rental Vacancy Rate</td>
<td>CENSUS</td>
<td>qly</td>
</tr>
</tbody>
</table>
Data used in this study:

- **Regional level**
  - 7 price series
  - 14 price and volume series.

- **National level**
  - 9 price series
  - 17 price and volume series.
Three Level Dynamic Factor model

For each series $i$ in region $b$:

$$X_{bit} = \Lambda_{G.bi}(L)G_{bt} + e_{Xbit}.$$ 

Let $G_t = (G_{1t} \ G_{2t} \ldots G_{Bt})'$ the set of regional factors.

Let $Y_t$ be observed aggregate indices:

$$
\begin{pmatrix}
G_t \\
Y_t
\end{pmatrix}
= \Lambda_F(L)F_t +
\begin{pmatrix}
e_{Gt} \\
e_{Yt}
\end{pmatrix}
$$

Identification:

- $\Lambda_{Gb}(0)$ is lower triangular with 1’s on the diagonal
- $\Lambda_F(0)$ is lower triangular with 1’s on the diagonal.
Figure 2

- **NAR**: Shows fluctuation over time with peaks and troughs.
- **CMHPI**: Similar pattern with notable spikes.
- **OFHEO**: Smaller fluctuations with less pronounced peaks.
- **Case–Shiller**: Smaller range with minor variations compared to others.

These charts illustrate the variability in different datasets over a range of time periods.
Figure 4: House Price vs. Housing Factor

- $F_p$
- $F_{pq}$
Conclusion

No model can serve all needs!
A multilevel state space model with unique features

- Hierarchy in factors of up to 4 levels
- Explicitly model within block correlations
- Multiple factors at each level
- Dynamic evolution of the factors modeled in an internally coherent fashion
- We provide algorithms for estimation and monitoring in a data rich environment
- Wide range of applications.
Thank You!