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 $b = 1, \ldots, B$, $i = 1, \ldots, N_b$,

$$\begin{aligned} X_{bit} &= \Lambda_{G.bi}(L) G_{bt} + e_{Xbit}, \\ \psi_{X.bi}(L) e_{Xbit} &= \epsilon_{Xbit}, \end{aligned}$$

A Three Level Model

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

$$
\begin{aligned}
X_{bit} &= \Lambda_{G.bi}(L) G_{bt} + e_{Xbit}, \\
\psi_{X.bi}(L) e_{Xbit} &= \epsilon_{Xbit},
\end{aligned}
$$

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

$$
\begin{aligned}
G_{bjt} &= \Lambda_{F.bj}(L) F_t + e_{Gbjt}, \\
\psi_{G.bj}(L) e_{Gbjt} &= \epsilon_{Gbjt}.
\end{aligned}
$$

A Three Level Model

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

$$\begin{aligned} X_{bit} &= \Lambda_{G.bi}(L) G_{bt} + e_{Xbit}, \\ \psi_{X.bi}(L) e_{Xbit} &= \epsilon_{Xbit}, \end{aligned}$$

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

$$\begin{aligned} G_{bjt} &= \Lambda_{F.bj}(L) F_t + e_{Gbjt}, \\ \psi_{G.bj}(L) e_{Gbjt} &= \epsilon_{Gbjt}. \end{aligned}$$

Level 3: For $r = 1, \ldots k_F$

$$\psi_{F.r}(L) F_{rt} = \epsilon_{Frt},$$

$\epsilon_{Xbit}, \epsilon_{Gbjt}, \epsilon_{Fkt} \sim iid(0, \sigma^2_{Xbi}, \sigma^2_{G_bj}, \sigma^2_{F_r})$.

A Four Level Model

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

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

with dyanmics

$$
\begin{aligned}
\psi_{Z.bsi}(L)e_{Zbsit} &= \epsilon_{Zbsit} \\
\Psi_{H.bs}(L)e_{Hbst} &= \epsilon_{Hbst} \\
\Psi_{G.b}(L)e_{Gbt} &= \epsilon_{Gbt}
\end{aligned}
$$

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

$$X_{bt} = \Lambda_{Gb.0} G_{bt} + \ldots + \Lambda_{Gb.s_{Gb}} G_{b,t-s_{Gb}} + e_{Xbt}$$

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:
e.g. $k_G = k_F = 1$:

$$\begin{aligned} x_{ibt} &= \lambda_{G.ib}(\lambda_{F.b} F_t + e_{G.bt}) + e_{X.ibt} \\ &= \lambda_{ib} F_t + v_{ibt} \\ v_{ibt} &= \lambda_{G.ib} e_{G.bt} + e_{X.ibt}. \end{aligned}$$

Ignoring block level variations gives a level 2 model:

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

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

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: $k_G << N$
- 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: $\widetilde{F}_t(\widetilde{G}_t)$:

- For each b, first estimate G_{bt}, then estimate F_t from $\widetilde{G}_t$
- 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:

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

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:

$$\begin{aligned} X_{bt} &= \Lambda_{Gb.0} G_{bt} + \ldots + \Lambda_{Gb.s_{Gb}} G_{b,t-s_{Gb}} + e_{Xbt} \\ \psi_{X.bi}(L) e_{Xbit} &= \epsilon_{Xbit}, \end{aligned}$$

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

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

Observed data:

$$\begin{aligned} X_{bt} &= \Lambda_{Gb.0} G_{bt} + \ldots + \Lambda_{Gb.s_{Gb}} G_{b,t-s_{Gb}} + e_{Xbt} \\ \psi_{X.bi}(L) e_{Xbit} &= \epsilon_{Xbit}, \end{aligned}$$

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

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

with time-varying intercept

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

Identification

- $var(\epsilon_{Gb})$ and $var(\epsilon_F)$ 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 $\boldsymbol{\Sigma} = (\Sigma_F, \Sigma_G, \Sigma_X)$, $\boldsymbol{\Psi} = (\Psi_F, \Psi_G, \Psi_X)$, $\boldsymbol{\Lambda} = (\Lambda_G, \Lambda_F)$

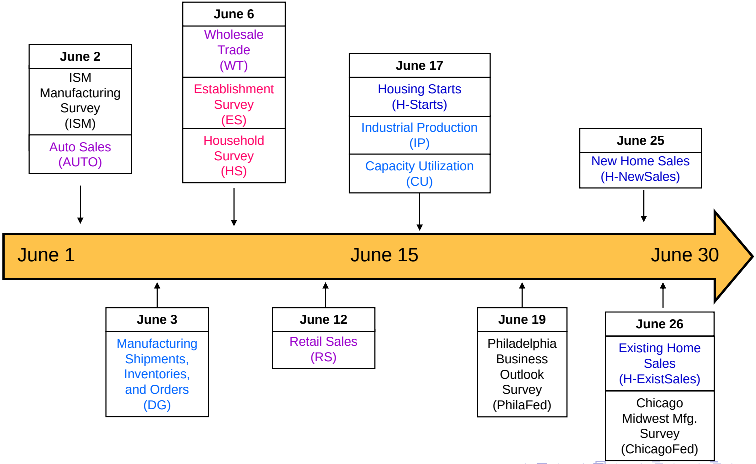
1. Use PCs as initial estimates of $\{G_t\}$ and $\{F_t\}$, get initial values for $\boldsymbol{\Lambda}$, $\boldsymbol{\Psi}$, $\boldsymbol{\Sigma}$
2. Conditional on $\boldsymbol{\Lambda}, \boldsymbol{\Psi}$, $\boldsymbol{\Sigma}$ and $\{F_t\}$: draw $\{G_{bt}\}$ block by block using Carter-Kohn algorithm modified to allow for time varying intercept
3. Conditional on $\boldsymbol{\Lambda}, \boldsymbol{\Psi}$, $\boldsymbol{\Sigma}$ and $\{G_t\}$: draw $\{F_t\}$
4. Conditional on $\{F_t\}$ and $\{G_t\}$: draw $\boldsymbol{\Lambda}$, $\boldsymbol{\Psi}$, and $\boldsymbol{\Sigma}$ assuming conjugate priors

Step 2: allows lower level factors to depend on the factors at the next level.
Level two transition equation: $\alpha_{F.bt} = \Psi_{Gb}(L)\Lambda_F(L)F_t$

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

- The time-varying intercept $\alpha_{F.bt}$ is known given $\mathbf{\Psi}$, $\mathbf{\Lambda}$, $\{F_t\}$:
- 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



A Three Level Model: 315 series

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

Parameters:

- $k_F = 1,\ k_G = 2$
- $s_F = 2,\ s_G = 2$
- $q_X = q_G = q_F = 1$

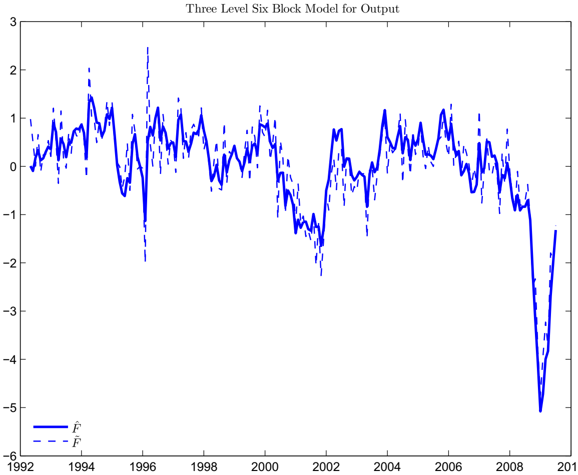
Table 4: Level 3, 6 Block Model

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

Decomposition of Variance

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

Figure: 6 Block Model of Real Activity:



Is $\widetilde{F}_t$ picking up block-level common variations?

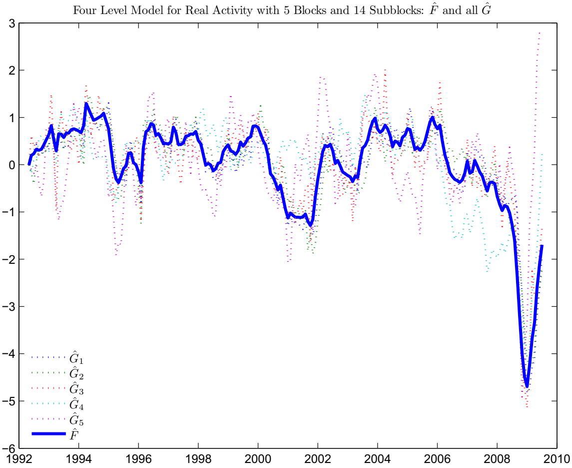
- Bai and Ng (2002): IC_2 finds 2 factors common to $\widetilde{G}_t$
- for $r = 1, \ldots k_F$, regress $\widetilde{F}_{rt}$ on $\widehat{F}_t$. Let $\widetilde{e}_{rt}$ be the residuals
- regress $\widetilde{e}_{rt}$ on $\widehat{G}_t$
- R^2 measures variations in $\widetilde{F}_{kt}$ that are not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

Table 5: Correlation Between $\widehat{G}_{bkt}$ and $\widetilde{e}_{rt|\widehat{F}}$

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

Four Level Model: 447 Series

Block	subblock	N	K_{Gb}	K_{Hbs}	Variable Ordered First
Production	CU	25	1	2	Capacity Utilization
	IP	38	1	2	IP: Durables
	DG	60	1	2	Manufacturers' Inventori
Employment	ES	82	1	2	All Employees: Wholesal
	HS	92	1	2	Civilian Labor Force: Me
Demand	RS	30	1	2	Retail Sales: General Me
	WS	54	1	2	Merchant Wholesalers: S
	AUTO	4	1	1	Domestic Car Retail Sale
Housing	H-STARTS	24	1	2	Housing Starts: 1-Unit:
	H-NEWSALES	5	1	1	New 1-Family Houses So
	H-EXISTSALES	4	1	1	NAR Total Existing Hom
Mfg Surveys	ISM	9	1	1	ISM Mfg: PMI Composit
	PHILAFED	21	1	1	Phila FRB Bus Outlook:
	CHICFED	5	1	1	Chicago FRB: Midwest M


Four Level Model for Real Activity with 5 Blocks and 14 Subblocks: $\hat{F}$ and all $\hat{G}$
3
2
1
0
−1
−2
−3
−4
−5
−6
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
$\hat{G}_1$
$\hat{G}_2$
$\hat{G}_3$
$\hat{G}_4$
$\hat{G}_5$
$\hat{F}$

Decomposition of Variance

block	sub-block	share_F	share_G	share_H	share_X
Output	IP	0.090	0.115	0.176	0.619
Output	CU	0.089	0.113	0.161	0.637
Output	DG	0.022	0.026	0.180	0.773
Employment	HS	0.042	0.090	0.226	0.642
Employment	ES	0.015	0.033	0.171	0.782
Demand	RS	0.030	0.073	0.212	0.685
Demand	WT	0.012	0.028	0.153	0.806
Demand	AUTO	0.018	0.049	0.393	0.539
Housing	H-starts	0.005	0.069	0.176	0.750
Housing	H-newsales	0.002	0.030	0.285	0.683
Housing	H-existsales	0.003	0.038	0.626	0.333
Mfg Surveys	ISM	0.035	0.161	0.248	0.556
Mfg Surveys	PhilaFed	0.010	0.045	0.197	0.747
Mfg Surveys	ChicagoFed	0.022	0.071	0.461	0.446

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

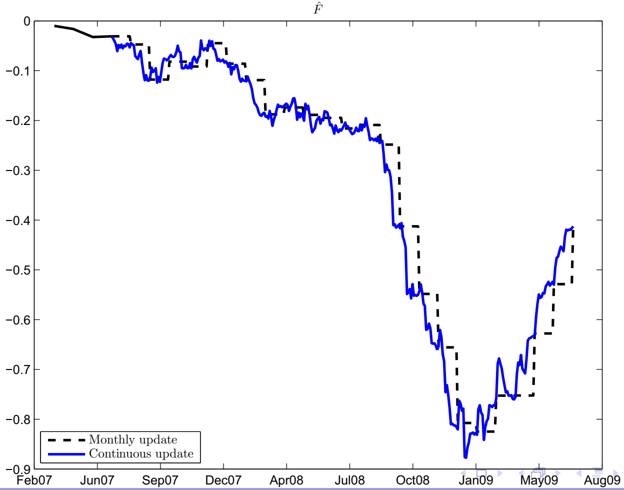


Figure: Demand Block

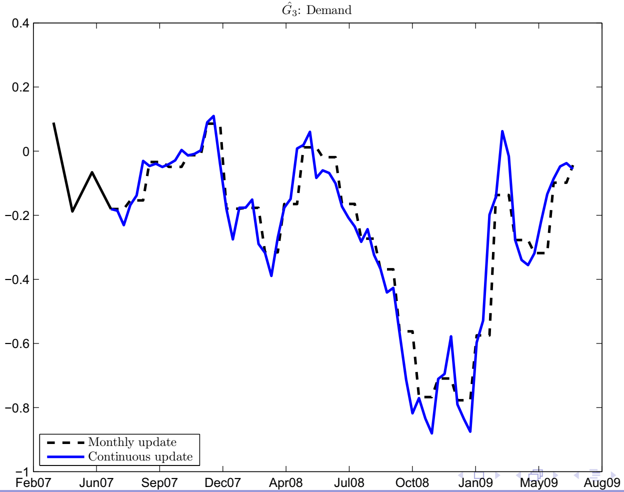


Figure: Housing

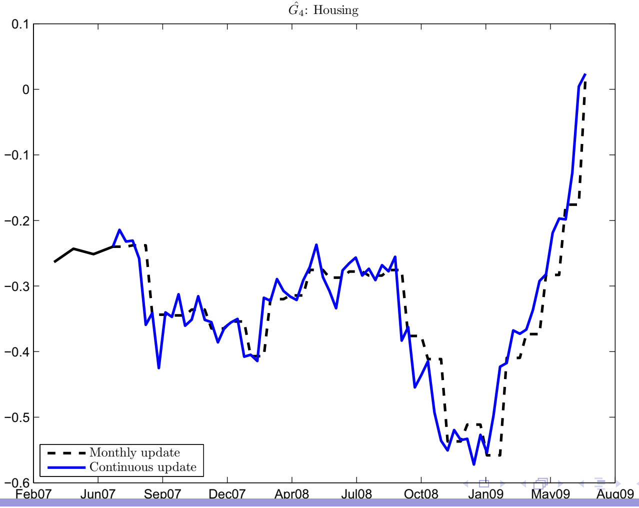
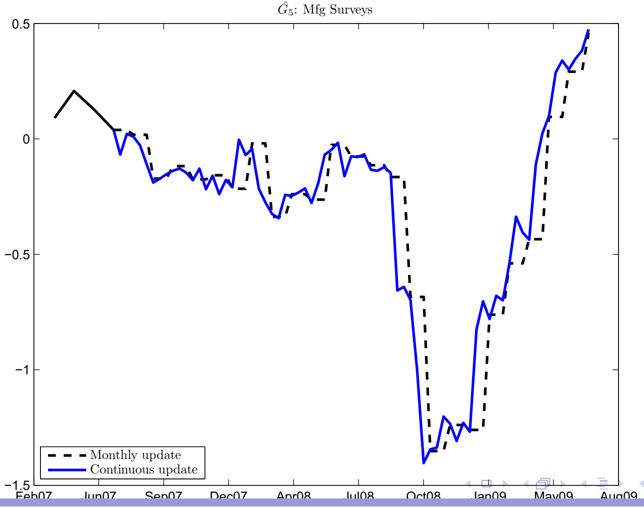


Figure: Manufacturing



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	Source	Frequency
Price data		
Median Sales Price of Single Family Existing Homes	NAR	mly
Single Family Median Home Sales Price	CENSUS	qly
Average Existing Home Prices	NAR	qly
Average New Home Prices	NAR	qly
Conventional Mortgage Home Price Index	FHLMC	qly
OFHEO Purchase-only Index	OFHEO	mly
OFHEO Home Prices	OFHEO	qly
Volume data		
New 1-Family Houses Sold	CENSUS	mly
New 1-Family Houses For Sale	CENSUS	mly
Single-Family Housing Units Under Construction	CENSUS	mly
Multifamily Units Under Construction	CENSUS	mly
Homeownership Rate	CENSUS	qly
Homeowner Vacancy Rate	CENSUS	qly
Rental Vacancy Rate	CENSUS	qly

National Series	Source	Frequency
Price data		
Median Sales Price of Single Family Existing Homes	NAR	mly
Median Sales Price of Single Family New Homes	Census	mly
Single Family Median Home Sales Price	CENSUS	qly
Average Existing Home Prices	NAR	mly
Average New Home Prices	CENSUS	mly
S&P/Case-Shiller Home Price Index	S&P	qly
Conventional Mortgage Home Price Index	FHLMC	qly
OFHEO Purchase-only Index	OFHEO	mly
OFHEO Home Prices	OFHEO	qly
Volume data		
New 1-Family Houses For Sale	CENSUS	mly
Housing Units Authorized by Permit: 1-Unit	CENSUS	mly
Multifamily Units Under Construction	CENSUS	mly
Multifamily Permits US	CENSUS	mly
Multifamily Starts US	CENSUS	mly
Multifamily Completions	CENSUS	mly

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

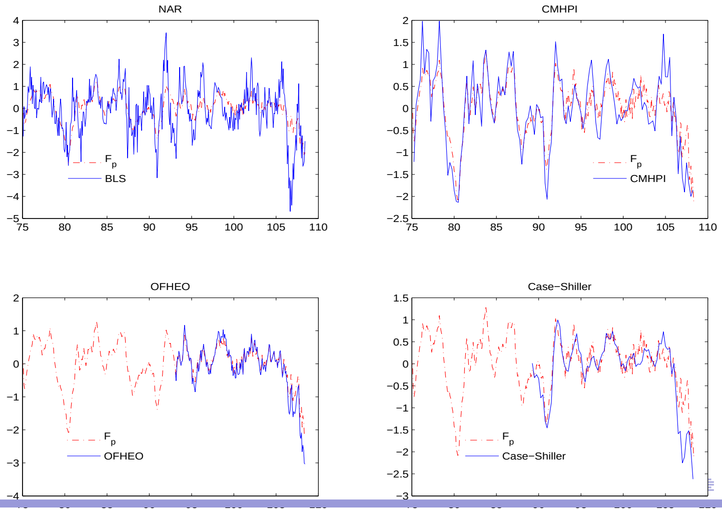
Let Y_t be observed aggregate indices:

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

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
Fp
BLS
CMHPI
Fp
CMHPI
OFHEO
Fp
OFHEO
Case−Shiller
Fp
Case−Shiller

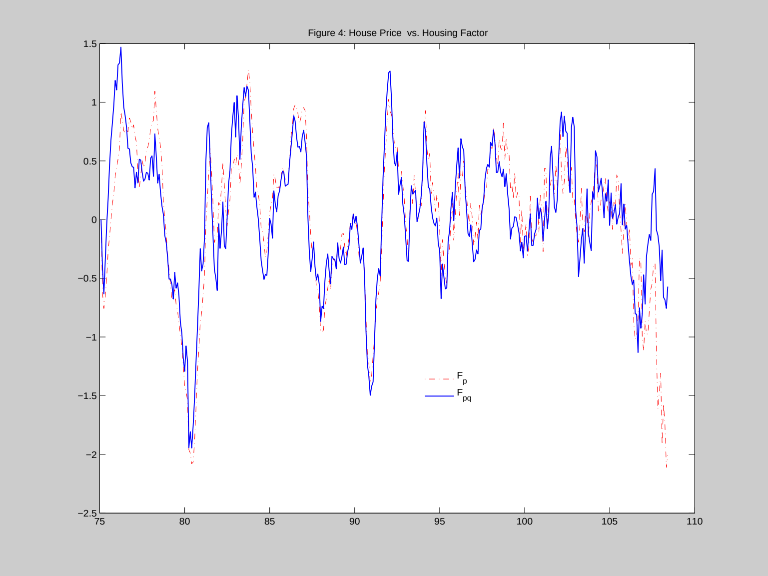


Figure 4: House Price vs. Housing Factor
1.5
1
0.5
0
−0.5
−1
−1.5
−2
−2.5
75
80
85
90
95
100
105
110
F_p
F_{pq}

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!