Dynamic Factor Models, Factor Augmented VARs, and SVARs in Macroeconomics

Mark Watson
Princeton University

Central Bank of Chile
October 22-24, 2018

Reference: Stock, James H. and Mark W. Watson (2016)
Handbook of Macroeconomics, Vol 2. chapter

Outline

Monday: Dynamic Factor Models – Part 1

Tuesday: Dynamic Factor Models – Part 2
SVARs – Part 1

Wednesday: SVARs – Part 2
FAVAR/SDFM

Historical Evolution of DFMs

I. Factor Analysis

- Spearman (1904)
- Lawley (1940), Joreskög (1967) … Lawley and Maxwell (1971)

Spearman's problem:

Data: X_{ij}, $i = 1, \ldots, N$ (individuals)

and $j = 1, \ldots n$ (measurements for each individual)

$$X_i = \begin{pmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{in} \end{pmatrix} \text{ and } \Sigma_{XX} = \text{cov}(X_i)$$

How can we measure 'intelligence'?

"GENERAL INTELLIGENCE," OBJECTIVELY DETERMINED AND MEASURED.

By C. SPEARMAN.

TABLE OF CONTENTS.

		PAGE
Chap. I.	Introductory.	
	1. Signs of Weakness in Experimental Psychology	202
	2. The Cause of this Weakness	203
	3. The Identities of Science	204
	4. Scope of the Present Experiments	205
Chap. II.	Historical and Critical	
	1. History of Previous Researches	206
	2. Conclusions to be drawn from these Previous Researches	219
	3. Criticism of Prevalent Working Methods	222
Chap. III.	Preliminary Investigation	
	1. Obviation of the Four Faults Quoted	225
	2. Definition of the Correspondence Sought	226
	3. Irrelevancies from Practice	227
	(a) Pitch	228
	(b) Sight	232
	(c) Weight	233
	(d) Intelligence	233
	4. Irrelevancies from Age	233
	5. Irrelevancies from Sex	235
	6. The Elimination of these Irrelevancies	236
	7. Alternations and Equivocalities	238
Chap. IV.	Description of the Present Experiments	
	1. Choice of Laboratory Psychics	241
	2. Instruments	242
	(a) Sound	243
	(b) Light	244
	(c) Weight	245
	3. Modes of Procedure	246
	(a) Experimental Series I	246
	(b) " " II	247
	(c) " " III	248
	(d) " " IV	249
	(e) " " V	249
	4. The Estimation of Intelligence	249
	5. Procedure in Deducing Results	252
	(a) Method of Correlation	252
	(b) Elimination of Observational Errors	253
	(c) Elimination of Irrelevant Factors	255
Chap. V.	The Present Results	
	1. Method and Meaning of Demonstration	256
	2. Correspondence between the Discrimination and the Intelligence	259

Experimental Series IV.

High Class Preparatory School for Boys.

A. Original Data.

Age		Pitch	Place in School (*before modification to eliminate Age*).												Music
			Classics			French			English			Mathem.			
Years	Months	Discrim. Thres. in 1/3 v. d., October, 1902	Xmas, 1902	Easter, 1903	July, 1903	Xmas, 1902	Easter, 1903	July, 1903	Xmas, 1902	Easter, 1903	July, 1903	Xmas, 1902	Easter, 1903	July, 1903	Ranked by Music Master
12	6	2	8	7	4	5	3	3	4	3	3	4	2	3	8
12	4	3	11	12	10	13	13	10	13	13	11	12	13	11	9
9	8	3	19	18	15	21	19	16	23	21	18	21	19	17	6
13	7	4	2	2	1	2	2	1	2	2	1	7	7	7	3
10	4	4	21		19	22		23	22		20	21		24	16
10	7	4	23	23	22	26	23	22	28	25	23	29	25	23	1
13	6	5	3			3			3			3			21
11	10	5	6	4	3	7	6	5	6	6	2	9	8	6	
10	1	5	29	26	24	23	25	21	27	26	22	25	23	19	7
11	1	6	20	20	18	20	21	18	21	20	19	17	16	15	14
13	4	7	1	1		1	1		1	1		1	1		5
10	6	7	26	24	21	27	16	13	26	19	17	22	18	16	11
12	3	7	18	17	16	17	20	19	25	23	21	19	17	14	20
13	1	8	5	5	5	4	4	2	5	8	5	5	4	1	4
11	1	10	22	19	17	19	18	17	20	17	15	23	21	21	18
9	9	10	33	29	27	33	29	27	33	27	27	32	29	27	17
10	4	11	28	25	23	30	27	24	18	18	13	30	27	22	
13	0	11	4	3	2	6	5	4	7	4	4	2	3	4	
10	2	11	7	6	6	12	7	6	8	5	8	11	9	8	
13	0	11	12	11	11	11	11	12	15	16	16	6	5	2	12
12	0	11	17	16		16	15		24	22		24	24		15
12	11	12	9	8	7	8	8	7	9	7	7	14	12	12	
13	1	14	10	9	8	10	9	8	11	10	9	10	10	9	13
10	4	14	27	21	14	24	22	15	17	11	10	26	20	18	2
10	1	15	24	22	20	18	17	14	29	24	24	18	15	13	
12	6	15	14	13	12	15	14	11	10	9	6	8	6	5	10
10	8	15	30	27		29	26		30	29		28	26		
12	8	18	16	15	13	25	24	20	14	14	12	20	21	20	19
9	5	20	32		25	31		25	32		26	33		26	
11	2	24	15	14	9	14	12	9	16	15	14	13	11	10	
10	9	50	25			28			19			15			
10	11	> 60	31	28	26	32	28	26	31	28	25	31	28	25	22
13	7	> 60	13	10		9	10		12	12		16	14		

Factor Model

$$X_{ij} = \lambda_j f_i + e_{ij} \text{ or}$$

$$X_i = \lambda f_i + e_i$$

$$\Sigma_{XX} = \sigma_f^2 \lambda\lambda' + \Sigma_{ee} \text{ with } \Sigma_{ee} \text{ diagonal}$$

$$X_i = \lambda f_i + e_i$$

$$\Sigma_{XX} = \sigma_f^2\ \lambda\lambda' + \Sigma_{ee} \text{ with } \Sigma_{ee} \text{ diagonal}$$

Issues:

(1) Estimation of parameters (σ_f^2, λ, $\sigma_{e_i}^2$) (Lawley: Gaussian MLE)

(2) Estimation of $f_i \mid X_i$, (σ_f^2, λ, $\sigma_{e_i}^2$): 'reverse regression'

$$(X_i \mid f_i) \sim \mathrm{N}(\lambda f_i, \Sigma_{ee}) \text{ and } f_i \sim \mathrm{N}(0, \sigma_f^2)$$

$$\Rightarrow f_i \mid X_i \sim \mathrm{N}(\beta' X_i\ ,\ \sigma_{f|Y}^2)$$

$$\text{with } \beta = \Sigma_{YY}^{-1}\Sigma_{Yf} = \left(\sigma_f^2 \lambda\lambda' + \Sigma_{ee}\right)^{-1} \lambda\sigma_f^2$$

$$\sigma_{f|Y}^2 = \sigma_f^2 - \sigma_f^2 \lambda' \left(\sigma_f^2 \lambda\lambda' + \Sigma_{ee}\right)^{-1} \lambda\sigma_f^2$$

Historical Evolution of DFMs:

2a: Replace covariance matrices with spectral density matrices. (Geweke (1977), Sargent and Sims (1977), Brillinger (1975)).

$$X_i = \lambda f_i + e_i$$

$$\Sigma_{XX} = \sigma_f^2 \lambda\lambda' + \Sigma_{ee} \text{ with } \Sigma_{ee} \text{ diagonal}$$

becomes

$$X_t = \lambda(\mathrm{L})f_t + e_t$$

$$S_{XX}(\omega) = s_f^2(\omega)\ \lambda(\mathrm{e}^{-\mathrm{i}\omega})\lambda(\mathrm{e}^{\mathrm{i}\omega})' + S_{ee}(\omega) \text{ with } S_{ee}(\omega) \text{ diagonal}$$

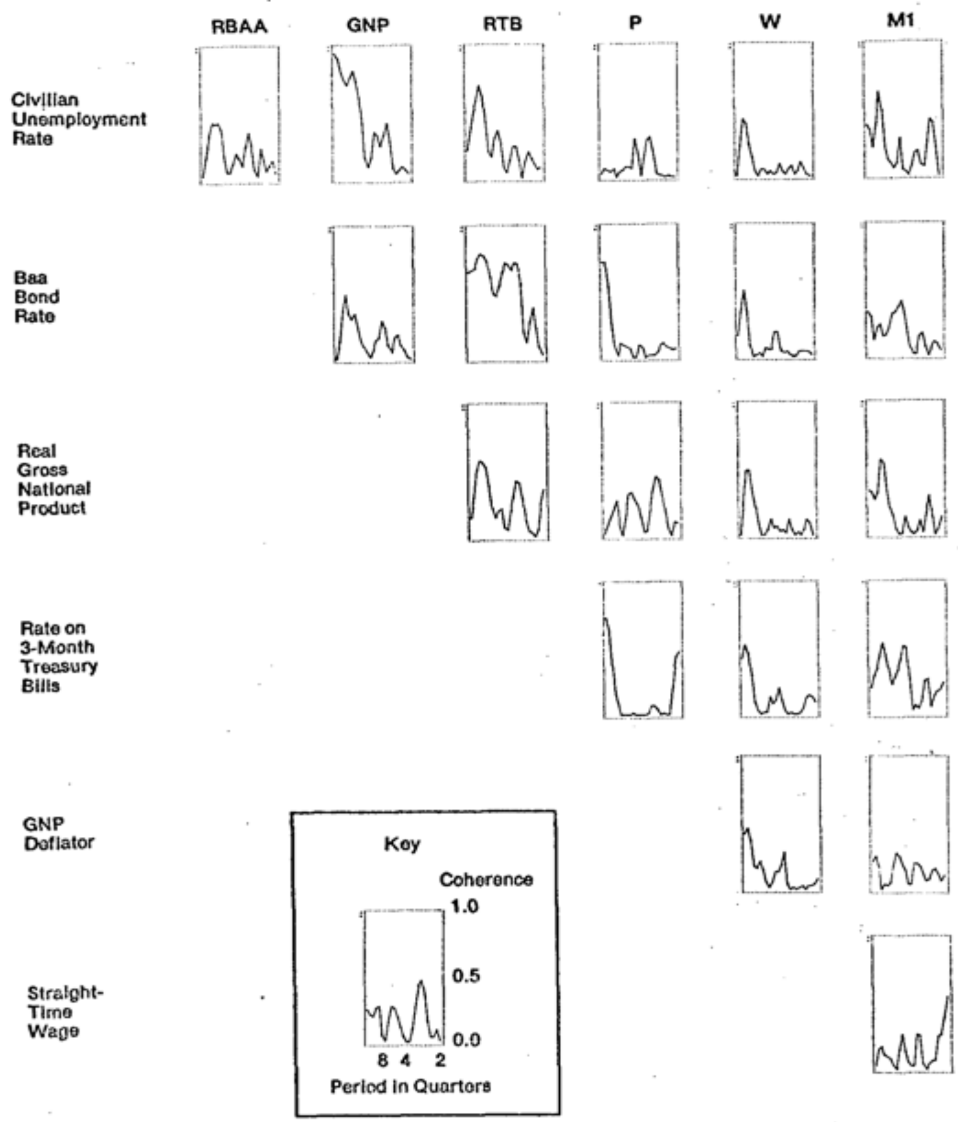
Business Cycle Modeling Without Pretending
to Have Too Much A Priori Economic Theory

Thomas J. Sargent
Christopher A. Sims

Revised, January 1977

Paper prepared for seminar on New Methods in Business Cycle Research, Federal Reserve Bank of Minneapolis, November 13-14, 1975. The views expressed herein are solely those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. John Geweke adapted the maximum likelihood factor analysis algorithm for application to the frequency domain factory model and wrote a computer program for estimating and testing the one-index model. Paul Anderson extended that program to handle k noises and performed all the frequency domain calculations in this paper. Salih Neftci carried out the calculations for the observable index model. John Geweke's contribution in developing the factor analysis algorithm and in formulating the unobservable index model were enough for him to qualify as a coauthor of this paper.

Table 1 — GRAPHS OF COHERENCE OF ECONOMIC VARIABLES



Sargent and Sims used various subsets of 14 variables: long rate, short rate, GNP, prices, wages, money supply, government purchases, government deficit, unemployment rate, residential construction, inventories, plant and equip investment, consumption, corporate profits.

$$X_t = \lambda(\mathrm{L})f_t + e_t$$

$$S_{XX}(\omega) = s_f^2(\omega)\ \lambda(\mathrm{e}^{-\mathrm{i}\omega})(\mathrm{e}^{\mathrm{i}\omega})\lambda' + S_{ee}(\omega) \text{ with } S_{ee}(\omega) \text{ diagonal}$$

Issues:

(1) Estimation of parameters ($s_f^2(\omega)$, $\lambda(\mathrm{e}^{-\mathrm{i}\omega})$, $S_{ee}(\omega)$) (Local Gaussian MLE, frequency by frequency)

(2) Estimation of $f(\omega) \mid X(\omega)$: can use 'reverse regression'

New issues: Converting frequency domain back to time domain. Leads/lags. Constraints across frequencies.

Table 12 Set 4

PROP OF VAR EXPLAINED BY 1 COMMON FACTORS

FREQUENCY	VAR. NO. 1 UNEMP.RT	VAR. NO. 2 LREALGNP	VAR. NO. 3 LGNPDEFL	VAR. NO. 4 LM1	VAR. NO. 5 LRES CONST	VAR. NO. 6 DLINVENT	VAR. NO. 7 LPL.EQPT	VAR. NO. 8 LCONS	VAR. NO. 9 LCORP.IVA
.1200PI	.84977	.93613	.21773	.28148	.26281	.54252	.68492	.12651	.99845
.3500PI	.43539	.91708	.97035E-01	.51548E-01	.80182E-01	.55580	.64545	.73225E-01	.87726
.6400PI	.55832	.73002	.24552	.73328E-01	.28865E-01	.21120	.28036	.10498E-01	.67750
.8600PI	.82680E-01	.25349	.29783	.33570E-01	.18589	.36609E-01	1.0000	.65040E-02	.25051
OVERALL	.82831	.93171	.21507	.27275	.21453	.53089	.67412	.85918E-01	.89032

PROP OF VAR EXPLAINED BY 2 COMMON FACTORS

FREQUENCY	VAR. NO. 1 UNEMP RT	VAR. NO. 2 LREALGNP	VAR. NO. 3 LGNPDEFL	VAR. NO. 4 LM1	VAR. NO. 5 LRES CONST	VAR. NO. 6 DLINVENT	VAR. NO. 7 LPL.EQPT	VAR. NO. 8 LCONS	VAR. NO. 9 LCORP.IVA
.1200PI	.97520	.93926	1.0000	.37270	.26820	.84758	.73196	.17193	.51112
.3500PI	.67835	1.0000	.43216	.79475E-01	.13698	.53886	.68603	.77118E-01	1.0000
.6400PI	.55357	.81562	.54733	.37229	.11344	.69574	1.0000	.29553E-01	.94345
.8600PI	.89870E-01	.77358	1.0000	.19129	.50045E-01	.51306	.56072	.735[illegible]	.86558
OVERALL	.86352	.93088	.98189	.36266	.22583	.83306	.77543	.11191	.91369

2b: Use linear state-space models: (e.g., Engle and Watson (1981))

$$X_t = \lambda(L)f_t + e_t \text{ and } \phi(L)f_t = \eta_t$$

$$X_t = \begin{pmatrix} \lambda_0 & \lambda_1 & \cdots & \lambda_k \end{pmatrix} \begin{pmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-k} \end{pmatrix} + e_t$$

$$\begin{pmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-k} \end{pmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{k+1} \\ 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} f_{t-1} \\ f_{t-2} \\ \vdots \\ f_{t-k-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \eta_t$$

or

$$X_t = \Lambda F_t + e_t$$
$$F_t = \Phi F_{t-1} + G\eta_t$$

(More generally *F* equation can be VAR(*p*))

Issues:

(1) Estimation of parameters (Λ, σ_η^2, Φ, Σ_{ee}) (Gaussian MLE using prediction-error decomposition from Kalman filter)

(2) Estimation of $f_t \mid \left\{X_j\right\}_{j=1}^{T}$: 'reverse regression' computed using Kalman smoother.

New issues:

(a) State-space modeling afforded lots of flexibility.

(b) MLE hard when X_t is high dimensional.

A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates

ROBERT ENGLE and MARK WATSON*

The paper formulates and estimates a single-factor multivariate time series model. The model is a dynamic generalization of the multiple indicator (or factor analysis) model. It is shown to be a special case of the general state space model and can be estimated by maximum likelihood methods using the Kalman filter algorithm. The model is used to obtain estimates of the unobserved metropolitan wage rate for Los Angeles, based on observations of sectoral wages within the Standard Metropolitan Statistical Area. Hypothesis tests, model diagnostics, and out-of-sample forecasts are used to evaluate the model.

KEY WORDS: State space model; Dynamic factor analysis; Kalman filter; Method of scoring; Unobserved component estimation.

1. INTRODUCTION

Much of the growth and decline of regional economies can be attributed to changes in comparative advantage, and the single most important component of this comparative advantage is probably wage rates. Therefore, considerable interest centers on the measurement of local wage rates and on the determinants of their movements. Because a region within a national economy can be thought of as a very open economy, there are strong economic pressures for wages to equalize between regions, both through commodity trade which tends to equate factor prices and through regional migration of labor and capital. For further discussion of these issues, see Engle (1974).

The measurement of a regional wage rate and its determinants is complicated by the differing wage in different industries and by differing skill mixes in different industries. In this article a statistical technique will be employed to separate movements in a metropolitan wage rate into a national industrial component, a metropolitan area-wide component, and a local industry specific component. For example, the wage rate in contract construction in Los Angeles will be decomposed into one component determined by the wage rate in contract construction in the United States, a second determined by the overall wage rate in Los Angeles, and a third resulting from factors particular to Los Angeles contract construction.

There are good economic reasons for expecting each of these components to be important. The national component measures not only changes in the U.S. economy as a whole through inflation and business cycles, but also measures changes in technology, changes in preferences, changes in the supply or demand for the output of the industry nationally, and collective bargaining outcomes that may affect industrial wages for a broad geographical region. The metropolitan component reflects the demand and supply of labor in the metropolitan labor market. Presumably, no industry can avoid the effect of the local labor market entirely, but some may be more strongly influenced than others. This component would reflect migration patterns of capital and labor, the cost of living in the region, and the tightness of the local labor market. The specific effect is the remainder which measures situations peculiar to this industry and region. By definition, the three effects are independent.

To illustrate the problem, consider the least squares regression of the log of the wage rate in industry i in Los Angeles, w_{it}, on the log of the national wage rate in this industry, n_{it}, using annual data. The residuals from this regression are composed of metropolitan effects and local industry specific effects. The metropolitan effects are common to each industry and therefore produce correlation across industries while the specific effects are by definition independent of other industries. In Table 1, these regressions and residual correlations are presented; the large cross-sectional correlations suggest the importance of the metropolitan effect. A factor analysis of these residual correlations indicates that one factor could explain 70 percent of the variance.

Because the data are a time series of cross-sections, the dynamic effects must also be considered and standard factor analysis is not appropriate. The first-order lagged correlation matrix, also presented in Table 1, shows the importance of the dynamics in the data set. Cross-correlations between sectors must result from serial correlation in the metropolitan component, while autocorrelations could arise from serial correlation in the specific effect. The frequency domain version of factor analysis of Geweke (1977) and Geweke and Singleton (1981) can

* Robert Engle is Professor, Department of Economics, University of California at San Diego, La Jolla, CA 92093. Mark Watson is Assistant Professor, Department of Economics, Harvard University, Cambridge, MA 02138. This research was supported by NSF grant SOC 77-07166. The authors are indebted to Clive W. J. Granger, David Lilien, Adrian Pagan, and Andrew Harvey for useful comments, suggestions, and encouragement at various stages of the research. The authors alone take credit for any remaining errors.

© Journal of the American Statistical Association
December 1981, Volume 76, Number 376
Applications Section

Table 3. Dynamic Factor Analysis (Model B)[a]

Where $m_t = \phi_1 m_{t-1} + \phi_2 m_{t-2} + v_{1t}$

$w_{it} = \alpha_i m_t + \beta_i n_{it} + e_{it}$ For sectors $i = 1, \ldots, 5$

$e_{it} = \rho_i e_{it-1} + v_{1+it}$

Sector	α	β	ρ	$\sigma^2 \times 10^4$	SE
Contract construction	1.	.874	.628	.598	.008
		(.078)	(.389)	(.329)	
Durable manufactures	.549	.786	.742	.835	.009
	(.090)	(.053)	(.155)	(.266)	
Nondurable manufactures	.380	.786	.898	.466	.007
	(.091)	(.040)	(.107)	(.149)	
Wholesale trade	.302	.959	.519	1.191	.011
	(.075)	(.032)	(.227)	(.352)	
Retail trade	.663	.810	.340	.941	.010
	(.070)	(.059)	(.289)	(.343)	

	ϕ_1	ϕ_2	$\sigma^2 \times 10^4$	σ_{v1}
Metropolitan component	1.606	−.619	1.229	.011
	(.125)	(.145)	(.585)	

[a] Standard errors are in parentheses.

Some Jargon:

$X_t = \lambda(L)f_t + e_t$ and $\phi(L)f_t = \eta_t$: *Dynamic form of DFM*

stacked version

$X_t = \Lambda F_t + u_t$ and $F_t = \Phi F_{t-1} + G\eta_t$: *Static form of DFM*

Example: "Improving GDP Measurement: A Measurement-Error Perspective" Aruoba, Diebold, Nalewaik, Schorfheide, Song (2016)

S.B. Aruoba et al. / Journal of Econometrics 191 (2016) 384–397 387

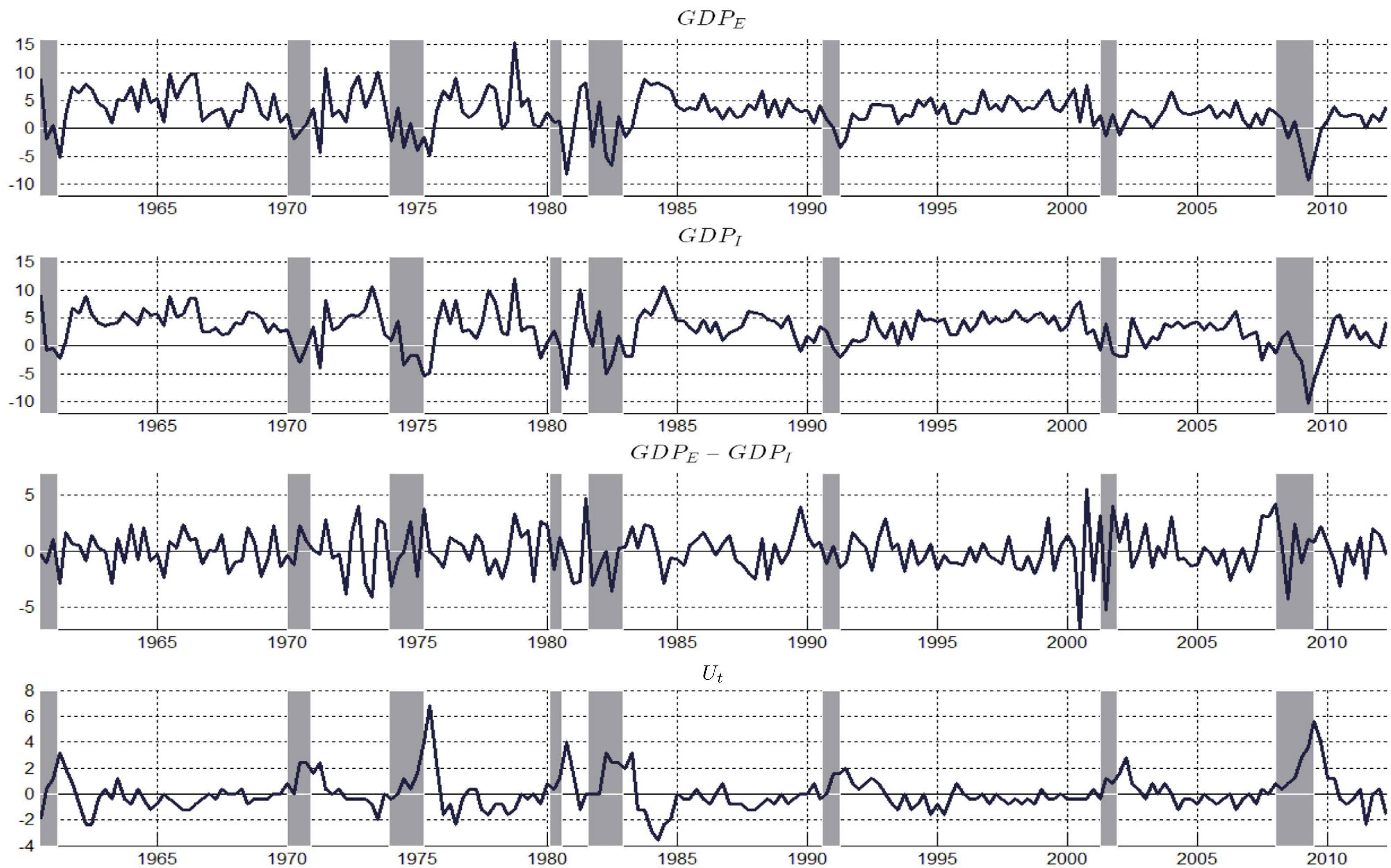


Fig. 1. *GDP* and unemployment data. GDP_E and GDP_I are in growth rates and U_t is in changes. All are measured in annualized percent.

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} GDP_t + \begin{bmatrix} \varepsilon_{Et} \\ \varepsilon_{It} \end{bmatrix}$$

$$GDP_t = \alpha + \rho GDP_{t-1} + \varepsilon_{Gt}$$

$$\text{var}\begin{bmatrix} \varepsilon_g \\ \varepsilon_E \\ \varepsilon_I \end{bmatrix} = \Sigma = \begin{bmatrix} \sigma_{GG} & 0 & 0 \\ & \sigma_{EE} & \sigma_{EI} \\ & & \sigma_{II} \end{bmatrix} \text{ (identification issues)}$$

Results:

For the 2-equation model with Σ block-diagonal, we have

$$GDP_t = \underset{[2.77,3.34]}{3.06}(1-0.62) + \underset{[0.57,0.68]}{0.62} GDP_{t-1} + \epsilon_{Gt}, \tag{12}$$

$$\Sigma = \begin{bmatrix} \underset{[4.39,5.95]}{5.17} & 0 & 0 \\ 0 & \underset{[3.34,4.48]}{3.86} & \underset{[0.96,1.95]}{1.43} \\ 0 & \underset{[0.96,1.95]}{1.43} & \underset{[2.25,3.22]}{2.70} \end{bmatrix}. \tag{13}$$

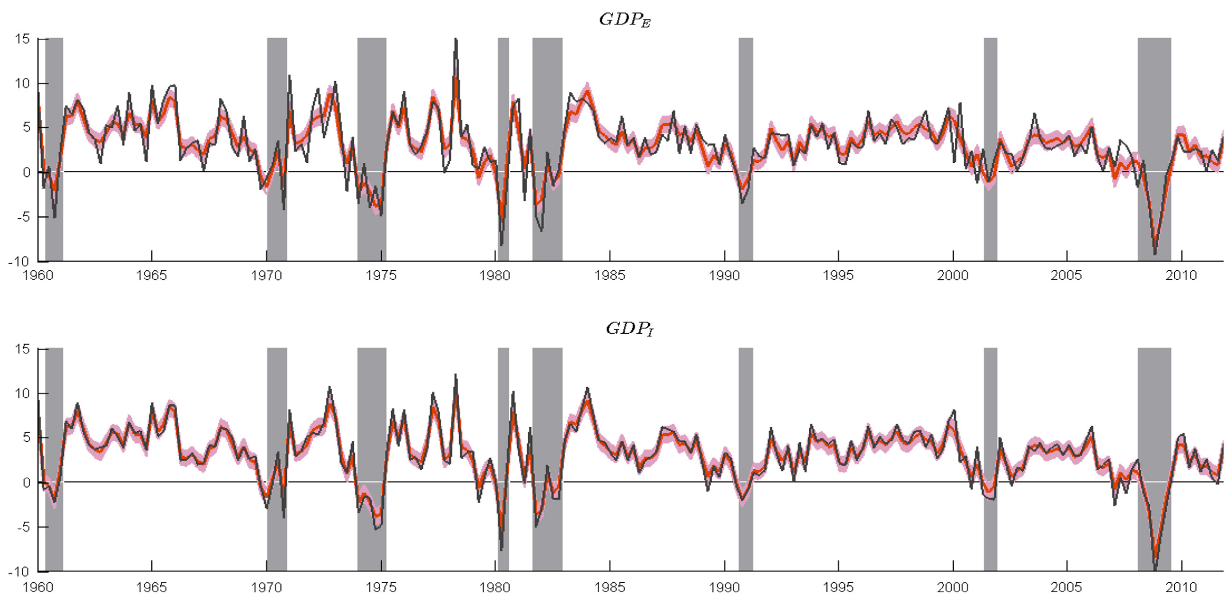
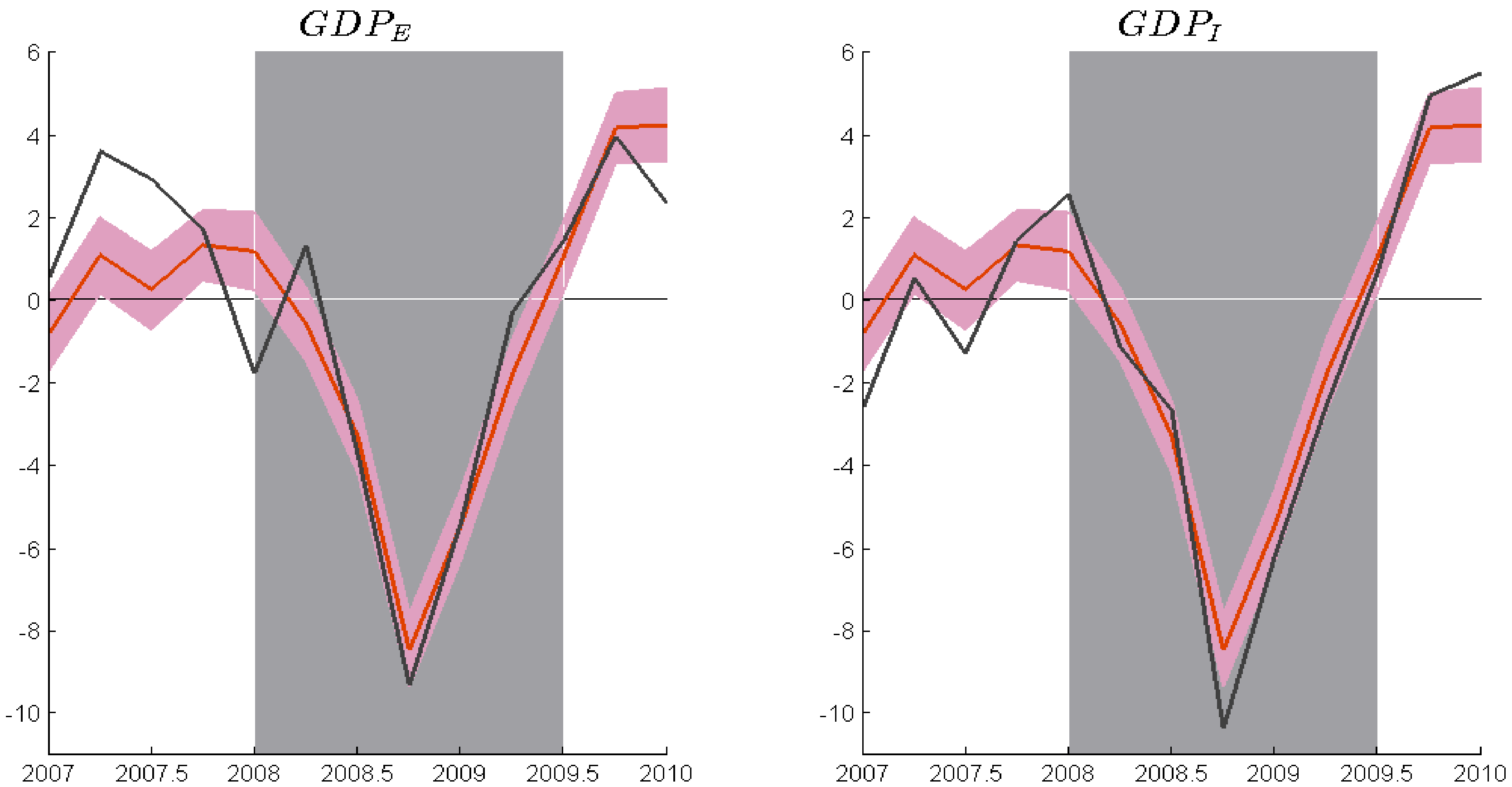


Fig. 3. *GDP* sample paths, 1960Q1–2011Q4. In each panel we show the sample path of GDP_M (light color) together with posterior interquartile range with shading and we show one of the competitor series (dark color). For GDP_M we use our benchmark estimate from the 2-equation model with $\zeta = 0.80$.

Figure 4: *GDP* Sample Paths, 2007Q1-2009Q4



Historical Evolution of DFMs:

3. Large-n approximations. Connor and Karijczyk (1986), Chamberlain and Rothschild (1983), Forni and Reichlin (1998), Stock and Watson (2002), …

Large n … from curse to blessing: *An example following Forni and Reichlin (1998)*. Suppose f_t is scalar and $\lambda(L) = \lambda$ ("no lags in the factor loadings"), so

$$X_{it} = \lambda_i f_t + e_{it} \quad \text{for } i = 1, \dots n$$

Then:

$$\frac{1}{n}\sum_{i=1}^{n} X_{it} = \frac{1}{n}\sum_{i=1}^{n}(\lambda_i f_t + e_{it}) = \left(\frac{1}{n}\sum_{i=1}^{n}\lambda_i\right) f_t + \frac{1}{n}\sum_{i=1}^{n} e_{it}$$

If the errors e_{it} have limited dependence across series, then as n gets large,

$$\frac{1}{n}\sum_{i=1}^{n} X_{it} \xrightarrow{p} \bar{\lambda} f_t$$

Large n lets us recover f_t up to a scale factor.

A "least squares" reason to use the sample mean.

Consider

$$\min_{\{f_t\},\{\lambda_i\}} \sum_{i,t} (X_{it} - \lambda_i f_t)^2 \quad \text{subject to } \bar{\lambda} = 1$$

Yields: $\hat{f}_t = \frac{1}{n}\sum_{i=1}^{n} X_{it}$

(Other normalizations: $T^{-1}\sum_{t=1}^{T} f_t^2 = 1$)

Multivariate Problem: $X_{it} = \lambda_i' F_t + e_{it}$, where λ_i' is i^{th} row of Λ.

$$\min_{\{f_t\},\{\lambda_i\}} \sum_{i,t} (X_{it} - \lambda_i' F_t)^2 \text{ subject } T^{-1} \sum_{t=1}^{T} F_t F_t' = \Gamma \text{ (diagonal, with } \gamma_i \geq \gamma_{i+1}\text{)}$$

Yields: $\hat{F}_t$ as the principal components (PC) of X_t, (i.e., the linear combinations of X_t with the largest variance).

Odds and ends:

- Missing data
- Weighted least squares
- …

More generally

$X_t = \lambda(L)f_t + e_t$ and $\phi(L)f_t = \eta_t \Rightarrow X_t = \Lambda F_t + e_t$ and $\Phi(L)F_t = G\eta_t$

So Principal Components (PC) can be used to estimate F in DFM.

A simple 2-step estimation problem:

(1) Estimate F_t by PC

(2) Estimate λ_i and var(e_{it}) from regression of X_{it} onto $\hat{F}_t$.

(3) Estimate dynamic equation for F using VAR with $\hat{F}_t$ replacing F.

Some results about these simple 2-step estimators when n and T are large:

Results for the exact static factor model:

Connor and Korajczyk (1986): consistency in the exact static FM with T fixed, $n \to \infty$.

Selected results for the approximate DFM: $X_t = \Lambda F_t + e_t$

Typical conditions (Stock-Watson (2002), Bai-Ng (2002, 2006)):

(a) $\frac{1}{T}\sum_{i=1}^{T} F_t F_t' \xrightarrow{p} \Sigma_F$ (stationary factors)

(b) $\Lambda'\Lambda/n \to$ (or $\xrightarrow{p}$) Σ_Λ Full rank factor loadings

(c) e_{it} are weakly dependent over time and across series

(d) F, e are uncorrelated at all leads and lags

Selected results for the approximate DFM, ctd.

Stock and Watson (2002a)

- consistency in the approximate DFM, n, $T \to \infty$.
- justify using $\hat{F}_t$ as a regressor (no errors-in-variable bias. etc.)
- oracle property for forecasts

Bai and Ng (2006)

- $N^2/T \to \infty$
- asymptotic normality of PC estimator of the common component at rate $\min(n^{1/2}, T^{1/2})$ in approximate DFM. These can be used to compute confidence sets for F_t.
- Similar results are rates for the two estimators of Λ, Φ, Σ_{ee} and $\Sigma_{\eta\eta}$.

Historical Evolution of DFMs:

An issue in PC estimates of DFMs: F_t is estimated using averages of X_t. This ignores information in leads and lags of X that would be utilized using optimal estimator (Kalman smoother).

4. Hybrid estimators: Use PCs to get first-round estimates of Λ, Φ, Σ_{ee} and $\Sigma_{\eta\eta}$, then use Kalman smoother to get estimates of F, or do MLE using these as initial guesses of parameters. (Doz, Giannone, Reichlin (2011, 2012).)

Example: Nowcasting (Good reference: Banbura, Giannoni, Modugno, and Reichlin (2013).)

- Problem: y_t is a variable of interest (e.g., GDP growth rate in quarter t). It is available with a lag (say in t+1 or t+2). X_t is a vector of variables that are measured *during* period t (and perhaps earlier). How do you guess the value of y_t given the X data that has been revealed.

- 'Solution': Suppose X_{t_1} denotes the information known at time t_1. Then best guess of y_t is $E(y_t | X_{t_1})$.
  - But how do you compute $E(y_t | X_{t_1})$?
  - How do you update the estimate as another element of X_t is revealed?

Giannone, Reichlin, et al modeling approach:

$$\begin{bmatrix} y_t \\ X_{1t} \\ \vdots \\ X_{nt} \end{bmatrix} = \begin{bmatrix} \lambda_y \\ \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} F_t + \begin{bmatrix} e_{yt} \\ e_{1t} \\ \vdots \\ e_{nt} \end{bmatrix}$$

$$\Phi(\mathrm{L})F_t = \eta_t$$

- $\mathrm{E}(y_t | X_{t_1}) = \lambda_y \times \mathrm{E}(F_t \mid X_{t_1})$
- $\mathrm{E}(F_t \mid X_{t_1})$ computed by Kalman filter

(Lots of details left out)

FEDERAL RESERVE BANK *of* NEW YORK *Serving the Second District and the Nation*

Nowcasting Report

We're sharing the MATLAB code for our nowcasting model on GitHub. Learn more on our blog.

OVERVIEW | NOWCAST | METHODOLOGY | FAQS

Oct 19, 2018: New York Fed Staff Nowcast

- The New York Fed Staff Nowcast stands at 2.1% for 2018:Q3 and 2.4% for 2018:Q4.

MORE

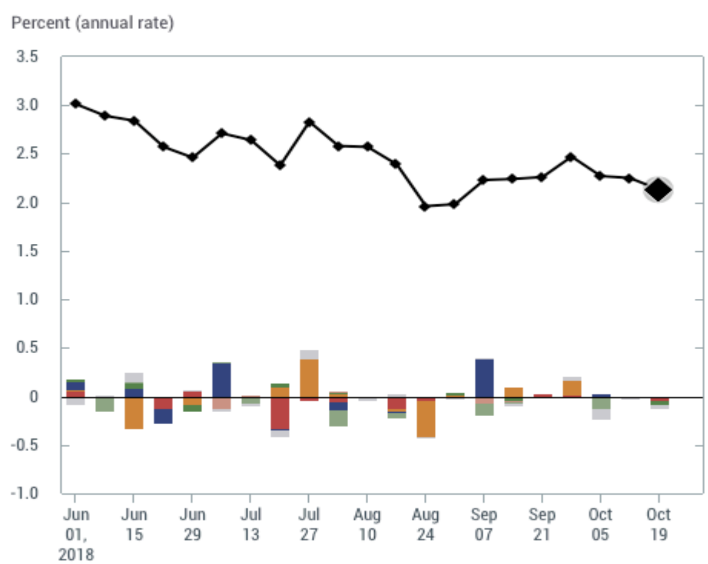
2018:Q4 | **2018:Q3** | 2018:Q2 | 2018:Q1

Last Release 11:15am EST Oct 19, 2018

ARCHIVE LAYOUT

◆ The New York Fed Staff Nowcast ○ Advance GDP estimate □ Latest GDP estimate

Housing and construction, Manufacturing, Surveys, Retail and consumption, Income, Labor, International trade, Others



Expand

Data Flow (Oct 19, 2018)

Model Update	Release Date	Data Series	Actual	Impact	Nowcast GDP Growth
Oct 19					2.13
	8:30AM Oct 18	Philadelphia Fed Mfg. Business Outlook: Current activity	22.20	0.00	
	8:30AM Oct 17	Building permits	-8.00	-0.01	
	8:30AM Oct 17	Housing starts	-5.28	-0.02	
	10:00AM Oct 16	JOLTS: Total job openings	59.00	-0.00	
	9:20AM Oct 16	Capacity utilization	0.05	-0.01	
	9:20AM Oct 16	Industrial production index	0.25	0.01	
	8:30AM Oct 15	Retail sales and food services	0.10	-0.05	
	8:30AM Oct 15	Empire State Mfg. Survey: General business conditions	21.10	0.00	
		Data revisions		-0.03	
Oct 12					2.25

Source: Authors' calculations, based on data accessed through Haver Analytics.

Notes: We start reporting the nowcast for a reference quarter about one month before the quarter begins; we stop updating it about one month after

2.1 | Nowcast Detail

■ Housing and construction ■ Manufacturing ■ Surveys ■ Retail and consumption ■ Income ■ Labor ■ International trade ■ Others

Update	Release Date	Data Series	Reference Period	Units	Forecast	Actual	Weight	Impact	Nowcast GDP Growth
					[a]	[b]	[c]	$[c(b-a)]$	
Sep 21									**2.66**
	10:00 AM Sep 26	New single family houses sold	Aug	MoM % chg.	0.236	3.45	0.008	0.025	
	8:30 AM Sep 27	Manufacturers' new orders: Durable goods	Aug	MoM % chg.	1.39	4.45	0.017	0.051	
	8:30 AM Sep 27	Merchant wholesalers: Inventories: Total	Aug	MoM % chg.	0.751	0.818	-0.078	-0.005	
	8:30 AM Sep 27	Manufacturers' shipments: Durable goods	Aug	MoM % chg.	0.730	0.753	0.106	0.002	
	8:30 AM Sep 27	Mfrs.' unfilled orders: All manufacturing industries	Aug	MoM % chg.	0.600	0.892	-0.008	-0.002	
	8:30 AM Sep 27	Manufacturers' inventories: Durable goods	Aug	MoM % chg.	0.693	-0.351	-0.185	0.194	
	8:30 AM Sep 28	Real disposable personal income	Aug	MoM % chg.	0.216	0.221	0.019	0.000	
	8:30 AM Sep 28	PCE less food and energy: Chain price index	Aug	MoM % chg.	0.157	0.037	0.293	-0.035	
	8:30 AM Sep 28	PCE: Chain price index	Aug	MoM % chg.	0.186	0.108	0.178	-0.014	
	8:40 AM Sep 28	Real personal consumption expenditures	Aug	MoM % chg.	0.229	0.222	0.270	-0.002	
		Data revisions						0.046	
Sep 28									**2.92**
	10:00 AM Oct 01	Value of construction put in place	Aug	MoM % chg.	0.379	0.080	0.027	-0.008	
	10:00 AM Oct 01	ISM mfg.: PMI composite index	Sep	Index	59.1	59.8	0.116	0.078	
	10:00 AM Oct 01	ISM mfg.: Prices index	Sep	Index	70.3	66.9	0.020	-0.068	
	10:00 AM Oct 01	ISM mfg.: Employment index	Sep	Index	57.3	58.8	0.053	0.079	
	8:05 AM Oct 03	ADP nonfarm private payroll employment	Sep	Level chg. (thousands)	177.4	229.0	1.358*	0.070	
	10:00 AM Oct 03	ISM nonmanufacturing: NMI composite index	Sep	Index	58.3	61.6	0.020	0.065	
	10:00 AM Oct 04	Inventories: Total business	Aug	MoM % chg.	0.397	0.479	-0.010	-0.001	
	8:30 AM Oct 05	All employees: Total nonfarm	Sep	Level chg. (thousands)	225.2	134.0	0.652*	-0.059	
	8:30 AM Oct 05	Civilian unemployment rate	Sep	Ppt. chg.	-0.084	-0.200	-0.186	0.022	
	8:30 AM Oct 05	Exports: Goods and services	Aug	MoM % chg.	1.30	-0.792	0.066	-0.138	
	8:30 AM Oct 05	Imports: Goods and services	Aug	MoM % chg.	1.00	0.586	0.050	-0.021	
		Data revisions						0.017	
		Parameter revisions						-0.153	
Oct 05									**2.80**
	8:30 AM Oct 10	PPI: Final demand	Sep	MoM % chg.	0.150	0.172	0.173	0.004	
	8:30 AM Oct 11	CPI-U: All items	Sep	MoM % chg.	0.229	0.059	0.151	-0.026	
	8:30 AM Oct 11	CPI-U: All items less food and energy	Sep	MoM % chg.	0.140	0.116	0.173	-0.004	
	8:30 AM Oct 12	Import price index	Sep	MoM % chg.	-0.039	0.470	0.045	0.023	
	8:30 AM Oct 12	Export price index	Sep	MoM % chg.	0.133	0.000	0.078	-0.010	
		Data revisions						-0.012	
Oct 12									**2.77**
	8:30 AM Oct 15	Empire State Mfg. Survey: General business conditions	Oct	Index	20.7	21.1	0.017	0.006	
	8:30 AM Oct 15	Retail sales and food services	Sep	MoM % chg.	0.560	0.104	0.353	-0.161	
	9:20 AM Oct 16	Industrial production index	Sep	MoM % chg.	0.182	0.252	0.399	0.028	
	9:20 AM Oct 16	Capacity utilization	Sep	Ppt. chg.	0.106	0.052	0.520	-0.028	
	10:00 AM Oct 16	JOLTS: Job openings: Total	Aug	Level chg. (thousands)	8.42	59.0	-0.048*	-0.002	
	8:30 AM Oct 17	Housing starts	Sep	MoM % chg.	-2.37	-5.28	0.018	-0.051	
	8:30 AM Oct 17	Building permits	Sep	Level chg. (thousands)	13.5	-8.00	0.002	-0.046	
	8:30 AM Oct 18	Phila. Fed Mfg. business outlook: Current activity	Oct	Index	26.4	22.2	0.012	-0.049	
		Data revisions						-0.042	
Oct 19									**2.43**

Source: Authors' calculations, based on data accessed through Haver Analytics.
Notes: MoM % chg. indicates month over month percentage change. QoQ % chg. indicates quarter over quarter percentage change. The weights with the asterisk are multiplied by 1,000 for legibility.

Historical Evolution of DFMs:

Issue: Many parameters in DFM. Shrinkage might be useful.

5. Bayes estimators (Kim and Nelson (1998), Otrok and Whiteman (1998))

$$X_t = \Lambda F_t + e_t \text{ and } \Phi(L)F_t = G\eta_t$$

Model is particularly amenable to MCMC methods:

(i) $(\Lambda, \Sigma_{ee}, \Phi, \Sigma_{\eta\eta} \mid \{X_t, F_t\})$: Linear regression problem

(ii) $(\{F_t\} \mid \{X_t\}, \Lambda, \Sigma_{ee}, \Phi, \Sigma_{\eta\eta})$: Linear signal extraction problem

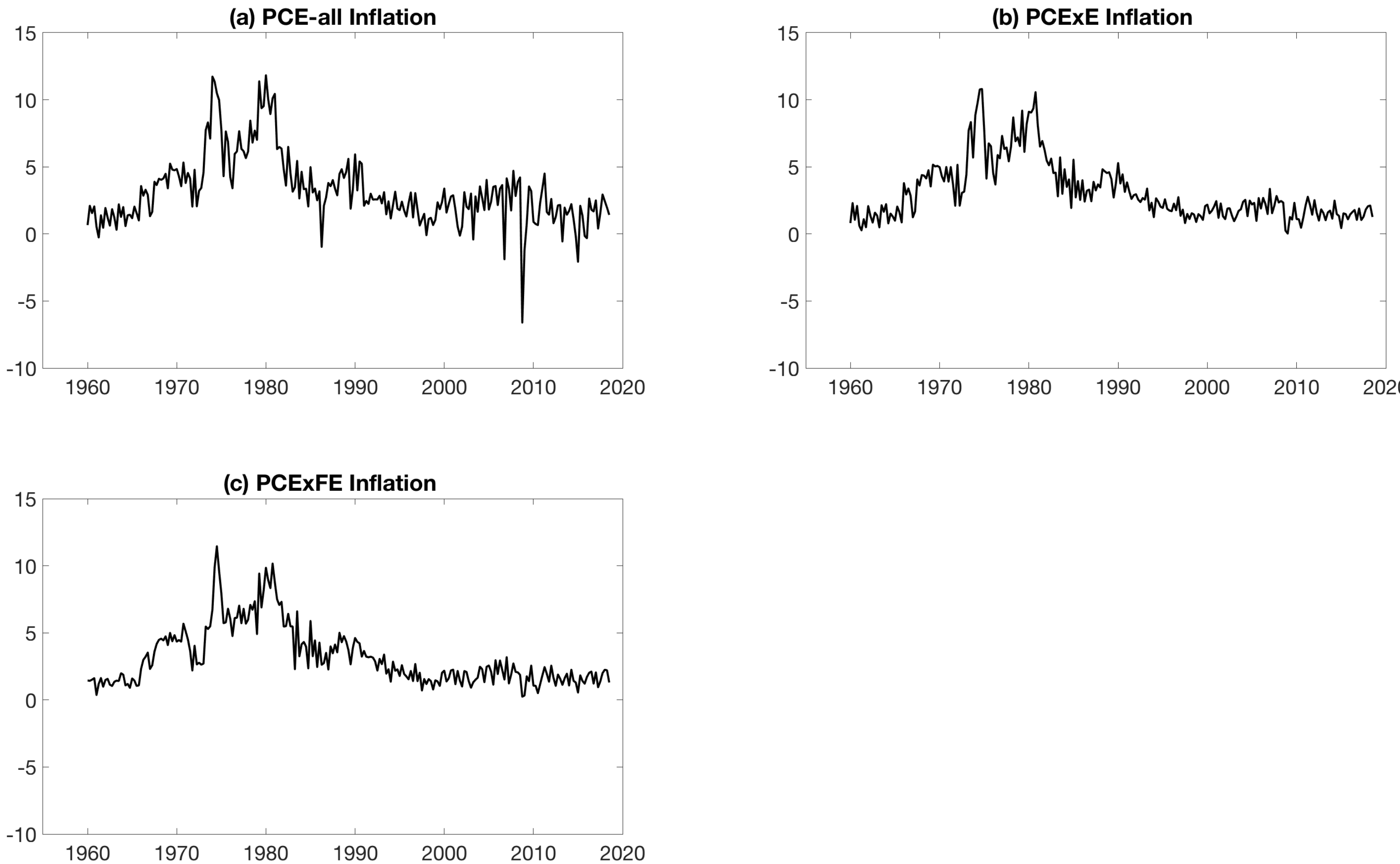
$$X_t = \Lambda F_t + e_t \text{ and } \Phi(\mathrm{L})F_t = \mathrm{G}\eta_t$$

Generalizations (see paper for references):

(1) Serial correlation in *e*
(2) Additional regressors in either equation
(3) Constraints on Λ ('sparsity')
(4) (Limited) cross-correlation between elements of *e*.
(5) Non-linearities and non-Gaussian evolution.

… many more.

Example (Non-linear and non-Gaussian): Stock and Watson (2016) 'Core Inflation and Trend Inflation' and earlier (2007) paper.



Unobserved Components Model with Stochastic Volatility and Outliers.

$$\boldsymbol{\pi_t = \tau_t + \varepsilon_t}$$

$$\tau_t = \tau_{t\text{-}1} + \sigma_{\Delta\tau,t} \times \eta_{\tau,t}$$

$$\varepsilon_t = \sigma_{\varepsilon,t} \times s_t \times \eta_{\varepsilon,t}$$

$$\Delta\ln(\sigma^2_{\varepsilon,t}) = \gamma_\varepsilon \nu_{\varepsilon,t}$$

$$\Delta\ln(\sigma^2_{\Delta\tau,t}) = \gamma_{\Delta\tau} \nu_{\Delta\tau,t}$$

$(\eta_\varepsilon, \eta_\tau, \nu_\varepsilon, \nu_{\Delta\tau})$ are iid $N(0, I_4)$

s_t = i.i.d. multinomial with values 1, 5, 10
and probability 0.975, 1/60, and 1/120

- Kim-Shephard-Chib (1998) approximate model for stochastic volatility:

Let $x_t = \sigma_t \eta_t$ and $\ln(\sigma_t^2) = \ln(\sigma_{t-1}^2) + \gamma \nu_t$ with $(\eta_t, \nu_t) \sim \text{iidN}(0, \text{I}_2)$.

Then $\ln(x_t^2) = \ln(\sigma_t^2) + \ln(\eta_t^2)$, where $\eta_t \sim \text{N}(0,1)$ so $\ln(\eta_t^2) \sim \ln(\chi_1^2)$

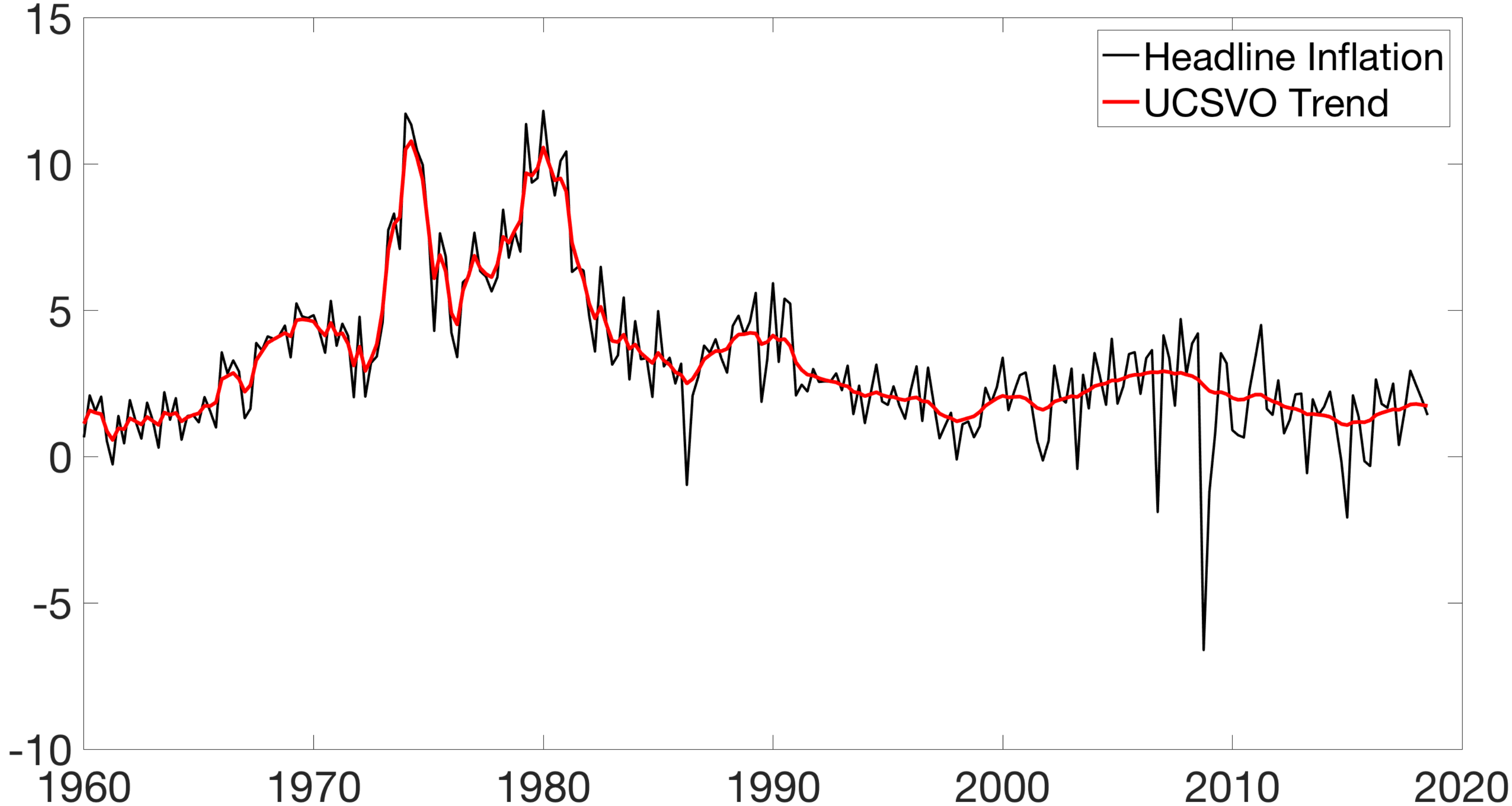
$\ln(\sigma_t^2) = \ln(\sigma_{t-1}^2) + \gamma \nu_t$

which is a linear state-space model with non-Gaussian measurement error.

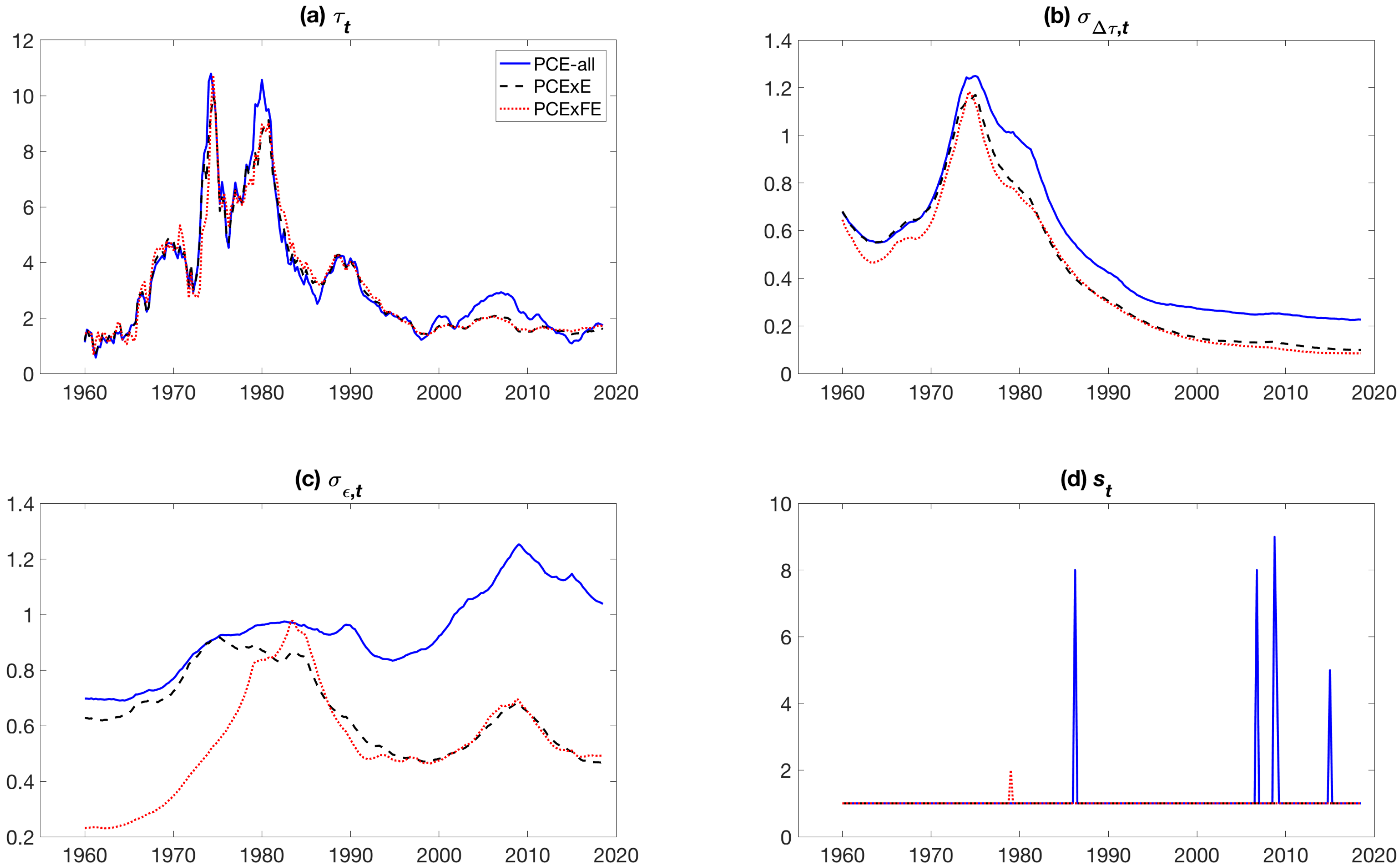
- KSC approximate $\ln(\chi_1^2)$ using a mixture of normals: $\ln(\eta_t^2) \sim \sum_{i=1}^{n} w_{it} a_{it}$,

where w_{it} are iid (0-1) variables with $w_{it} = 1$ for only value of i at each t, and with $p(w_{it} = 1) = p_i$. The a_{it} variables are $a_{it} \sim N(\mu_i, \sigma_i^2)$, and $n = 7$.

○ Omori, Chib, Shephard, and Nakajima (2007) propose a more accurate 10-component Gaussian mixture approximation.

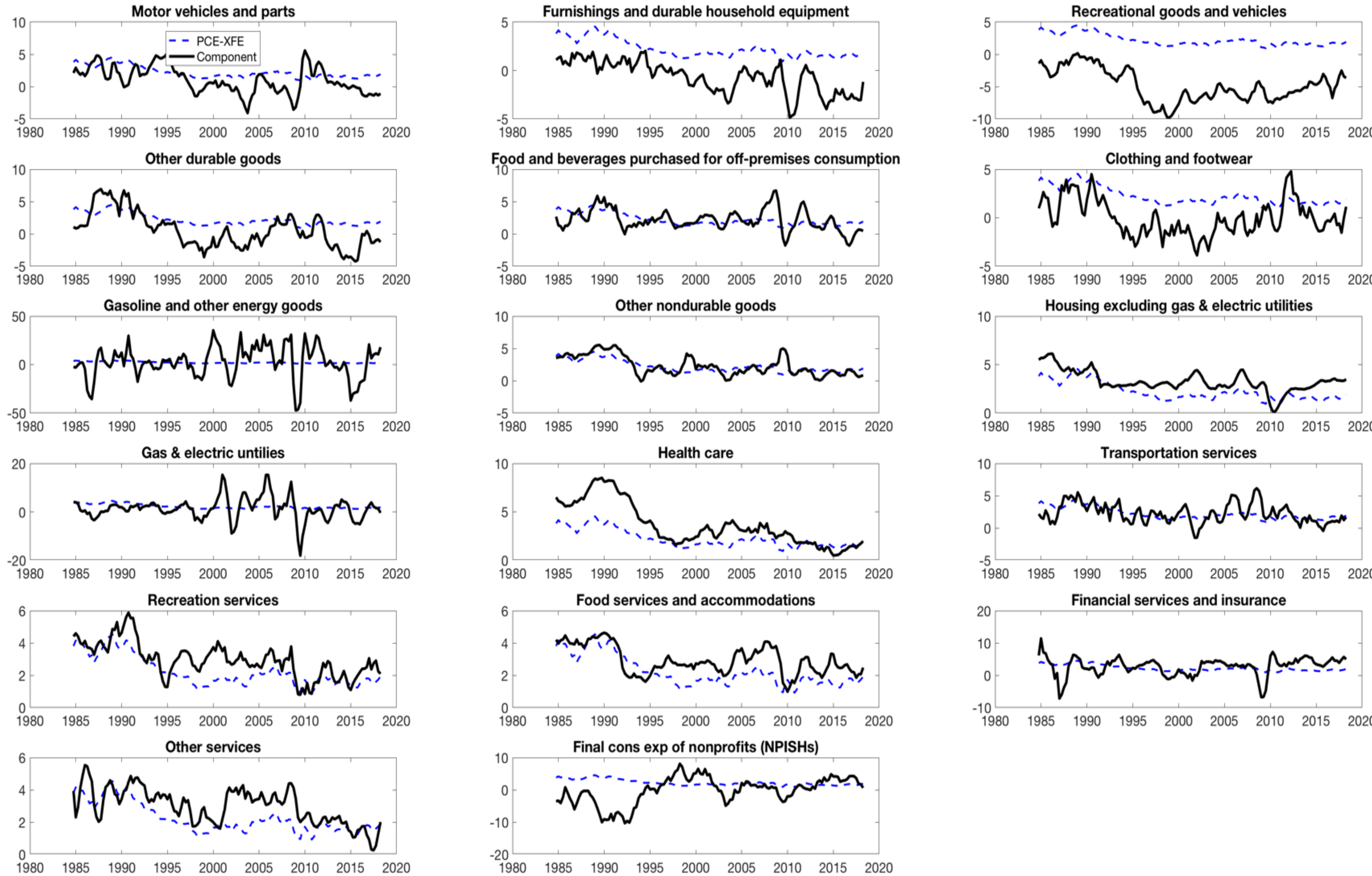


15
10
5
0
-5
-10
1960
1970
1980
1990
2000
2010
2020
Headline Inflation
UCSVO Trend


(a) τ_t
PCE-all
PCExE
PCExFE
(b) $\sigma_{\Delta\tau,t}$
(c) $\sigma_{\epsilon,t}$
(d) s_t
1960
1970
1980
1990
2000
2010
2020

17 PCE Sectors

Sector	Share
Motor vehicles and parts	0.042
Furnishings and durable household equip.	0.027
Recreational goods and vehicles	0.031
Other durable goods	0.016
Food and bev.s purch. for off-premises cons.*	0.077
Clothing and footwear	0.033
Gasoline and other energy goods*	0.030
Other nondurable goods	0.081
Housing & utilities	0.182
Housing excluding gas & electric utilities	0.162
Gas & electric utilities*	0.020
Health care	0.158
Transportation services	0.033
Recreation services	0.039
Food services and accommodations	0.063
Financial services and insurance	0.076
Other services	0.085
Final cons exp of nonprof. insti. serving h.h.	0.028


Motor vehicles and parts
PCE-XFE
Component
Furnishings and durable household equipment
Recreational goods and vehicles
Other durable goods
Food and beverages purchased for off-premises consumption
Clothing and footwear
Gasoline and other energy goods
Other nondurable goods
Housing excluding gas & electric utilities
Gas & electric untilies
Health care
Transportation services
Recreation services
Food services and accommodations
Financial services and insurance
Other services
Final cons exp of nonprofits (NPISHs)

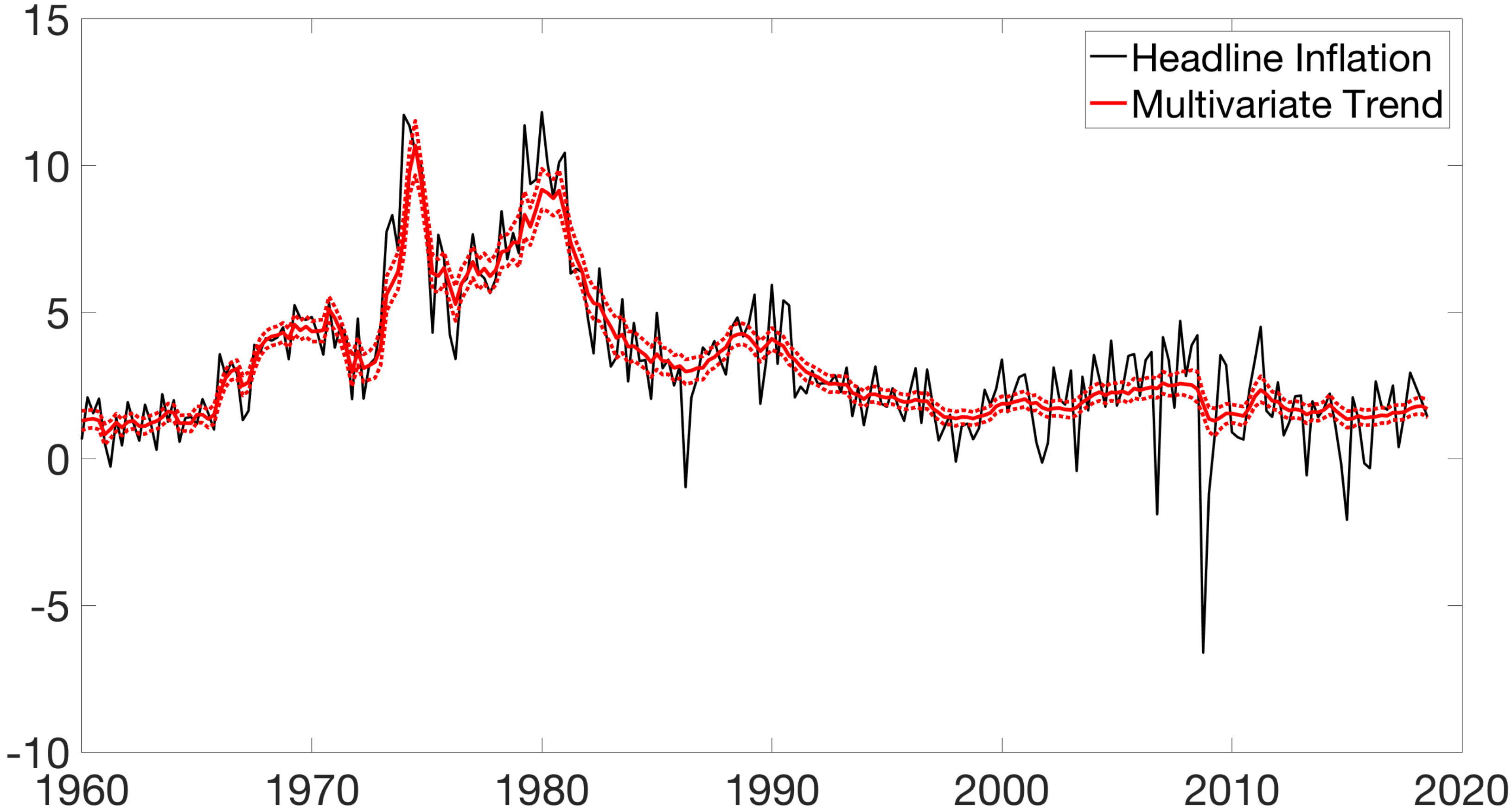
Multivariate model

$$\begin{bmatrix} \pi_{1t} \\ \pi_{2t} \\ \vdots \\ \pi_{nt} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}_t \tau_t^C + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \varepsilon_t^C + \begin{bmatrix} \tau_{1t}^u \\ \tau_{2t}^u \\ \vdots \\ \tau_{2t}^u \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t}^u \\ \varepsilon_{2t}^u \\ \vdots \\ \varepsilon_{2t}^u \end{bmatrix}$$

Aggregate (average) inflation and trend

$$\begin{aligned} \bar{\pi}_t &= \left[\bar{\alpha}\tau_t^c + \bar{\tau}_t^u \right] + \left[\bar{\beta}\varepsilon_t^c + \bar{\varepsilon}_t^u \right] \\ &= \qquad \tau_t \qquad + \qquad \varepsilon_t \end{aligned}$$

where the averages are computed using consumption share weights.


Headline Inflation
Multivariate Trend
15
10
5
0
-5
-10
1960
1970
1980
1990
2000
2010
2020

Recent Values of Inflation in the United States
(Quarterly inflation in percentage points at an annual rate)

	Inflation measures			Estimates from 17 component model	
Date	Headline	XFE		Trend	67% Band
2016:Q3	1.81	2.11		1.49	1.29 – 1.70
2016:Q4	1.67	1.21		1.48	1.29 – 1.68
2017:Q1	2.50	2.05		1.60	1.40 – 1.80
2017:Q2	0.40	0.94		1.49	1.28 – 1.70
2017:Q3	1.56	1.38		1.52	1.32 – 1.73
2017:Q4	2.94	2.00		1.62	1.41 – 1.83
2018:Q1	2.45	2.25		1.70	1.48 – 1.92
2018:Q2	1.96	2.21		1.82	1.57 – 2.06
2018:Q3	1.42	1.30		1.69	1.42 – 1.96

A 207-Variable Macro Dataset for the U.S.

Table 1 Quarterly time series in the full dataset

	Category	**Number of series**	**Number of series used for factor estimation**
(1)	NIPA	20	12
(2)	Industrial production	11	7
(3)	Employment and unemployment	45	30
(4)	Orders, inventories, and sales	10	9
(5)	Housing starts and permits	8	6
(6)	Prices	37	24
(7)	Productivity and labor earnings	10	5
(8)	Interest rates	18	10
(9)	Money and credit	12	6
(10)	International	9	9
(11)	Asset prices, wealth, and household balance sheets	15	10
(12)	Other	2	2
(13)	Oil market variables	10	9
	Total	207	139

Notes: The real activity dataset consists of the variables in the categories 1–4.

Table A.1: Data Series

	Name	Description	Sample Period	T	O	F
	(1) NIPA					
1	GDP	Real Gross Domestic Product 3 Decimal	1959:Q1-2014:Q4	5	0	0
2	Consumption	Real Personal Consumption Expenditures	1959:Q1-2014:Q4	5	0	0
3	Cons:Dur	Real Personal Consumption Expenditures: Durable Goods Quantity Index	1959:Q1-2014:Q4	5	0	1
4	Cons:Svc	Real Personal Consumption Expenditures: Services Quantity Index	1959:Q1-2014:Q4	5	0	1
5	Cons:NonDur	Real Personal Consumption Expenditures: Nondurable Goods Quantity Index	1959:Q1-2014:Q4	5	0	1
6	Investment	Real Gross Private Domestic Investment 3 Decimal	1959:Q1-2014:Q4	5	0	0
7	FixedInv	Real Private Fixed Investment Quantity Index	1959:Q1-2014:Q4	5	0	0
8	Inv:Equip	Real Nonresidential Investment: Equipment Quantity Idenx	1959:Q1-2014:Q4	5	0	1
9	FixInv:NonRes	Real Private Nonresidential Fixed Investment Quantity Index	1959:Q1-2014:Q4	5	0	1
10	FixedInv:Res	Real Private Residential Fixed Investment Quantity Index	1959:Q1-2014:Q4	5	0	1
11	Ch. Inv/GDP	Change in Inventories /GDP	1959:Q1-2014:Q4	1	0	1
12	Gov.Spending	Real Government Consumption Expenditures & Gross Investment 3 Decimal	1959:Q1-2014:Q4	5	0	0
13	Gov:Fed	Real Federal Consumption Expenditures Quantity Index	1959:Q1-2014:Q4	5	0	1
14	Real_Gov Receipts	Government Current Receipts (Nominal) Defl by GDP Deflator	1959:Q1-2014:Q3	5	0	1
15	Gov:State&Local	Real State & Local Consumption Expenditures Quantity Index	1959:Q1-2014:Q4	5	0	1
16	Exports	Real Exports of Goods & Services 3 Decimal	1959:Q1-2014:Q4	5	0	1
17	Imports	Real Imports of Goods & Services 3 Decimal	1959:Q1-2014:Q4	5	0	1
18	Disp-Income	Real Disposable Personal Income	1959:Q1-2014:Q4	5	0	0
19	Ouput:NFB	Nonfarm Business Sector: Output	1959:Q1-2014:Q4	5	0	0
20	Output:Bus	Business Sector: Output	1959:Q1-2014:Q4	5	0	0
	(2) Industrial Production					
21	IP: Total index	IP: Total index	1959:Q1-2014:Q4	5	0	0
22	IP: Final products	Industrial Production: Final Products (Market Group)	1959:Q1-2014:Q4	5	0	0
23	IP: Consumer goods	IP: Consumer goods	1959:Q1-2014:Q4	5	0	0
24	IP: Materials	Industrial Production: Materials	1959:Q1-2014:Q4	5	0	0
25	IP: Dur gds materials	Industrial Production: Durable Materials	1959:Q1-2014:Q4	5	0	1
26	IP: Nondur gds materials	Industrial Production: nondurable Materials	1959:Q1-2014:Q4	5	0	1
27	IP: Dur Cons. Goods	Industrial Production: Durable Consumer Goods	1959:Q1-2014:Q4	5	0	1
28	IP: Auto	IP: Automotive products	1959:Q1-2014:Q4	5	0	1
29	IP:NonDur Cons God	Industrial Production: Nondurable Consumer Goods	1959:Q1-2014:Q4	5	0	1
30	IP: Bus Equip	Industrial Production: Business Equipment	1959:Q1-2014:Q4	5	0	1
31	Capu Tot	Capacity Utilization: Total Industry	1967:Q1-2014:Q4	1	0	1
	(3) Employment and Unemployment					
32	Emp:Nonfarm	Total Nonfarm Payrolls: All Employees	1959:Q1-2014:Q4	5	0	0
33	Emp: Private	All Employees: Total Private Industries	1959:Q1-2014:Q4	5	0	0

34	Emp: mfg	All Employees: Manufacturing	1959:Q1-2014:Q4	5	0	0
35	Emp:Services	All Employees: Service-Providing Industries	1959:Q1-2014:Q4	5	0	0
36	Emp:Goods	All Employees: Goods-Producing Industries	1959:Q1-2014:Q4	5	0	0
37	Emp: DurGoods	All Employees: Durable Goods Manufacturing	1959:Q1-2014:Q4	5	0	1
38	Emp: Nondur Goods	All Employees: Nondurable Goods Manufacturing	1959:Q1-2014:Q4	5	0	0
39	Emp: Const	All Employees: Construction	1959:Q1-2014:Q4	5	0	1
40	Emp: Edu&Health	All Employees: Education & Health Services	1959:Q1-2014:Q4	5	0	1
41	Emp: Finance	All Employees: Financial Activities	1959:Q1-2014:Q4	5	0	1
42	Emp: Infor	All Employees: Information Services	1959:Q1-2014:Q4	5	1	1
43	Emp: Bus Serv	All Employees: Professional & Business Services	1959:Q1-2014:Q4	5	0	1
44	Emp:Leisure	All Employees: Leisure & Hospitality	1959:Q1-2014:Q4	5	0	1
45	Emp:OtherSvcs	All Employees: Other Services	1959:Q1-2014:Q4	5	0	1
46	Emp: Mining/NatRes	All Employees: Natural Resources & Mining	1959:Q1-2014:Q4	5	1	1
47	Emp:Trade&Trans	All Employees: Trade Transportation & Utilities	1959:Q1-2014:Q4	5	0	1
48	Emp: Gov	All Employees: Government	1959:Q1-2014:Q4	5	0	0
49	Emp:Retail	All Employees: Retail Trade	1959:Q1-2014:Q4	5	0	1
50	Emp:Wholesal	All Employees: Wholesale Trade	1959:Q1-2014:Q4	5	0	1
51	Emp: Gov(Fed)	Employment Federal Government	1959:Q1-2014:Q4	5	2	1
52	Emp: Gov (State)	Employment State government	1959:Q1-2014:Q4	5	0	1
53	Emp: Gov (Local)	Employment Local government	1959:Q1-2014:Q4	5	0	1
54	Emp: Total (HHSurve)	Emp Total (Household Survey)	1959:Q1-2014:Q4	5	0	0
55	LF Part Rate	LaborForce Participation Rate (16 Over) SA	1959:Q1-2014:Q4	2	0	0
56	Unemp Rate	Urate	1959:Q1-2014:Q4	2	0	0
57	Urate_ST	Urate Short Term (< 27 weeks)	1959:Q1-2014:Q4	2	0	0
58	Urate_LT	Urate Long Term (>= 27 weeks)	1959:Q1-2014:Q4	2	0	0
59	Urate: Age16-19	Unemployment Rate - 16-19 yrs	1959:Q1-2014:Q4	2	0	1
60	Urate:Age>20 Men	Unemployment Rate - 20 yrs. & over Men	1959:Q1-2014:Q4	2	0	1
61	Urate: Age>20 Women	Unemployment Rate - 20 yrs. & over Women	1959:Q1-2014:Q4	2	0	1
62	U: Dur<5wks	Number Unemployed for Less than 5 Weeks	1959:Q1-2014:Q4	5	0	1
63	U:Dur5-14wks	Number Unemployed for 5-14 Weeks	1959:Q1-2014:Q4	5	0	1
64	U:dur>15-26wks	Civilians Unemployed for 15-26 Weeks	1959:Q1-2014:Q4	5	0	1
65	U: Dur>27wks	Number Unemployed for 27 Weeks & over	1959:Q1-2014:Q4	5	0	1
66	U: Job losers	Unemployment Level - Job Losers	1967:Q1-2014:Q4	5	0	1
67	U: LF Reenty	Unemployment Level - Reentrants to Labor Force	1967:Q1-2014:Q4	5	1	1
68	U: Job Leavers	Unemployment Level - Job Leavers	1967:Q1-2014:Q4	5	0	1
69	U: New Entrants	Unemployment Level - New Entrants	1967:Q1-2014:Q4	5	1	1
70	Emp:SlackWk	Employment Level - Part-Time for Economic Reasons All Industries	1959:Q1-2014:Q4	5	1	1
71	EmpHrs:Bus Sec	Business Sector: Hours of All Persons	1959:Q1-2014:Q4	5	0	0
72	EmpHrs:nfb	Nonfarm Business Sector: Hours of All Persons	1959:Q1-2014:Q4	5	0	0
73	AWH Man	Average Weekly Hours: Manufacturing	1959:Q1-2014:Q4	1	0	1
74	AWH Privat	Average Weekly Hours: Total Private Industry	1964:Q1-2014:Q4	2	0	1
75	AWH Overtime	Average Weekly Hours: Overtime: Manufacturing	1959:Q1-2014:Q4	2	0	1
76	HelpWnted	Index of Help-Wanted Advertising in Newspapers (Data truncated in 2000)	1959:Q1-1999:Q4	1	0	0

	(4) Orders, Inventories, and Sales					
77	MT Sales	Manufacturing and trade sales (mil. Chain 2005 $)	1959:Q1-2014:Q3	5	0	0
78	Ret. Sale	Sales of retail stores (mil. Chain 2000 $)	1959:Q1-2014:Q3	5	0	1
79	Orders (DurMfg)	Mfrs' new orders durable goods industries (bil. chain 2000 $)	1959:Q1-2014:Q4	5	0	1
80	Orders (Cons. Gds & Mat.)	Mfrs' new orders consumer goods and materials (mil. 1982 $)	1959:Q1-2014:Q4	5	0	1
81	UnfOrders(DurGds)	Mfrs' unfilled orders durable goods indus. (bil. chain 2000 $)	1959:Q1-2014:Q4	5	0	1
82	Orders(NonDefCap)	Mfrs' new orders nondefense capital goods (mil. 1982 $)	1959:Q1-2014:Q4	5	0	1
83	VendPerf	ISM Manufacturing: Supplier Deliveries Index©	1959:Q1-2014:Q4	1	0	1
84	NAPM:INV	ISM Manufacturing: Inventories Index©	1959:Q1-2014:Q4	1	0	1
85	NAPM:ORD	ISM Manufacturing: New Orders Index©; Index;	1959:Q1-2014:Q4	1	0	1
86	MT Invent	Manufacturing and trade inventories (bil. Chain 2005 $)	1959:Q1-2014:Q3	5	0	1
	(5) Housing Starts and Permits					
87	Hstarts	Housing Starts: Total: New Privately Owned Housing Units Started	1959:Q1-2014:Q3	5	0	0
88	Hstarts >5units	Privately Owned Housing Starts: 5-Unit Structures or More	1959:Q1-2014:Q3	5	0	0
89	Hpermits	New Private Housing Units Authorized by Building Permit	1960:Q1-2014:Q4	5	0	1
90	Hstarts:MW	Housing Starts in Midwest Census Region	1959:Q1-2014:Q3	5	0	1
91	Hstarts:NE	Housing Starts in Northeast Census Region	1959:Q1-2014:Q3	5	0	1
92	Hstarts:S	Housing Starts in South Census Region	1959:Q1-2014:Q3	5	0	1
93	Hstarts:W	Housing Starts in West Census Region	1959:Q1-2014:Q3	5	0	1
94	Constr. Contracts	Construction contracts (mil. sq. ft.) (Copyright McGraw-Hill)	1963:Q1-2014:Q4	4	0	1
	(6) Prices					
95	PCED	Personal Consumption Expenditures: Chain-type Price Index	1959:Q1-2014:Q4	6	0	0
96	PCED_LFE	Personal Consumption Expenditures: Chain-type Price Index Less Food and Energy	1959:Q1-2014:Q4	6	0	0
97	GDP Defl	Gross Domestic Product: Chain-type Price Index	1959:Q1-2014:Q4	6	0	0
98	GPDI Defl	Gross Private Domestic Investment: Chain-type Price Index	1959:Q1-2014:Q4	6	0	1
99	BusSec Defl	Business Sector: Implicit Price Deflator	1959:Q1-2014:Q4	6	0	1
100	PCED_Goods	Goods	1959:Q1-2014:Q4	6	0	0
101	PCED_DurGoods	Durable goods	1959:Q1-2014:Q4	6	0	0
102	PCED_NDurGoods	Nondurable goods	1959:Q1-2014:Q4	6	0	0
103	PCED_Serv	Services	1959:Q1-2014:Q4	6	0	0
104	PCED_HouseholdServices	Household consumption expenditures (for services)	1959:Q1-2014:Q4	6	0	0
105	PCED_MotorVec	Motor vehicles and parts	1959:Q1-2014:Q4	6	0	1
106	PCED_DurHousehold	Furnishings and durable household equipment	1959:Q1-2014:Q4	6	0	1
107	PCED_Recreation	Recreational goods and vehicles	1959:Q1-2014:Q4	6	0	1
108	PCED_OthDurGds	Other durable goods	1959:Q1-2014:Q4	6	0	1
109	PCED_Food_Bev	Food and beverages purchased for off-premises consumption	1959:Q1-2014:Q4	6	0	1
110	PCED_Clothing	Clothing and footwear	1959:Q1-2014:Q4	6	0	1
111	PCED_Gas_Enrgy	Gasoline and other energy goods	1959:Q1-2014:Q4	6	0	1
112	PCED_OthNDurGds	Other nondurable goods	1959:Q1-2014:Q4	6	0	1

113	PCED_Housing-Utilities	Housing and utilities	1959:Q1-2014:Q4	6	0	1
114	PCED_HealthCare	Health care	1959:Q1-2014:Q4	6	0	1
115	PCED_TransSvg	Transportation services	1959:Q1-2014:Q4	6	0	1
116	PCED_RecServices	Recreation services	1959:Q1-2014:Q4	6	0	1
117	PCED_FoodServ_Acc.	Food services and accommodations	1959:Q1-2014:Q4	6	0	1
118	PCED_FIRE	Financial services and insurance	1959:Q1-2014:Q4	6	0	1
119	PCED_OtherServices	Other services	1959:Q1-2014:Q4	6	0	1
120	CPI	Consumer Price Index For All Urban Consumers: All Items	1959:Q1-2014:Q4	6	0	0
121	CPI_LFE	Consumer Price Index for All Urban Consumers: All Items Less Food & Energy	1959:Q1-2014:Q4	6	0	0
122	PPI:FinGds	Producer Price Index: Finished Goods	1959:Q1-2014:Q4	6	0	0
123	PPI	Producer Price Index: All Commodities	1959:Q1-2014:Q3	6	0	0
124	PPI:FinConsGds	Producer Price Index: Finished Consumer Goods	1959:Q1-2014:Q4	6	0	1
125	PPI:FinConsGds (Food)	Producer Price Index: Finished Consumer Foods	1959:Q1-2014:Q4	6	0	1
126	PPI:IndCom	Producer Price Index: Industrial Commodities	1959:Q1-2014:Q4	6	0	1
127	PPI:IntMat	Producer Price Index: Intermediate Materials: Supplies & Components	1959:Q1-2014:Q4	6	0	1
128	Real_P:SensMat	Index of Sensitive Matrerials Prices (Discontinued) Defl by PCE(LFE) Def	1959:Q1-2004:Q1	5	0	1
129	Real_Commod: spot price	Spot market price index:BLS & CRB: all commodities(1967=100) Defl by PCE(LFE)	1959:Q1-2009:Q1	5	0	0
130	NAPM com price	ISM Manufacturing: Prices Paid Index©	1959:Q1-2014:Q4	1	0	1
131	Real_Price:NatGas	PPI: Natural Gas Defl by PCE(LFE)	1967:Q1-2014:Q4	5	0	1
	(7) Productivity and Earnings					
132	Real_AHE:PrivInd	Average Hourly Earnings: Total Private Industries Defl by PCE(LFE)	1964:Q1-2014:Q4	5	0	0
133	Real_AHE:Const	Average Hourly Earnings: Construction Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	0
134	Real_AHE:MFG	Average Hourly Earnings: Manufacturing Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	0
135	CPH:NFB	Nonfarm Business Sector: Real Compensation Per Hour	1959:Q1-2014:Q4	5	0	1
136	CPH:Bus	Business Sector: Real Compensation Per Hour	1959:Q1-2014:Q4	5	0	1
137	OPH:nfb	Nonfarm Business Sector: Output Per Hour of All Persons	1959:Q1-2014:Q4	5	0	1
138	OPH:Bus	Business Sector: Output Per Hour of All Persons	1959:Q1-2014:Q4	5	0	0
139	ULC:Bus	Business Sector: Unit Labor Cost	1959:Q1-2014:Q4	5	0	0
140	ULC:NFB	Nonfarm Business Sector: Unit Labor Cost	1959:Q1-2014:Q4	5	0	1
141	UNLPay:nfb	Nonfarm Business Sector: Unit Nonlabor Payments	1959:Q1-2014:Q4	5	0	1
	(8) Interest Rates					
142	FedFunds	Effective Federal Funds Rate	1959:Q1-2014:Q4	2	0	1
143	TB-3Mth	3-Month Treasury Bill: Secondary Market Rate	1959:Q1-2014:Q4	2	0	1
144	TM-6MTH	6-Month Treasury Bill: Secondary Market Rate	1959:Q1-2014:Q4	2	0	0
145	EuroDol3M	3-Month Eurodollar Deposit Rate (London)	1971:Q1-2014:Q4	2	0	0
146	TB-1YR	1-Year Treasury Constant Maturity Rate	1959:Q1-2014:Q4	2	0	0
147	TB-10YR	10-Year Treasury Constant Maturity Rate	1959:Q1-2014:Q4	2	0	0
148	Mort-30Yr	30-Year Conventional Mortgage Rate	1971:Q2-2014:Q4	2	0	0
149	AAA Bond	Moody's Seasoned Aaa Corporate Bond Yield	1959:Q1-2014:Q4	2	0	0
150	BAA Bond	Moody's Seasoned Baa Corporate Bond Yield	1959:Q1-2014:Q4	2	0	0

151	BAA_GS10	BAA-GS10 Spread	1959:Q1-2014:Q4	1	0	1
152	MRTG_GS10	Mortg-GS10 Spread	1971:Q2-2014:Q4	1	0	1
153	tb6m_tb3m	tb6m-tb3m	1959:Q1-2014:Q4	1	0	1
154	GS1_tb3m	GS1_Tb3m	1959:Q1-2014:Q4	1	0	1
155	GS10_tb3m	GS10_Tb3m	1959:Q1-2014:Q4	1	0	1
156	CP_Tbill Spread	CP3FM-TB3MS	1959:Q1-2014:Q4	1	0	1
157	Ted_spr	MED3-TB3MS (Version of TED Spread)	1971:Q1-2014:Q4	1	0	1
158	gz_spread	Gilchrist-Zakrajsek Spread (Unadjusted)	1973:Q1-2012:Q4	1	0	0
159	gz_ebp	Gilchrist-Zakrajsek Excess Bond Premium	1973:Q1-2012:Q4	1	0	1
	(9) Money and Credit					
160	Real_mbase	St. Louis Adjusted Monetary Base; Bil. of $; M; SA; Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	0
161	Real_InsMMF	Institutional Money Funds Defl by PCE(LFE)	1980:Q1-2014:Q4	5	0	0
162	Real_m1	M1 Money Stock Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	0
163	Real_m2	M2SL Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	0
164	Real_mzm	MZM Money Stock Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	0
165	Real_C&Lloand	Commercial and Industrial Loans at All Commercial Banks Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	1
166	Real_ConsLoans	Consumer (Individual) Loans at All Commercial Banks/ Outlier Code because of change in data in April 2010. See FRB H8 Release Defl by PCE(LFE)	1959:Q1-2014:Q4	5	1	1
167	Real_NonRevCredit	Total Nonrevolving Credit Outstanding Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	1
168	Real_LoansRealEst	Real Estate Loans at All Commercial Banks Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	1
169	Real_RevolvCredit	Total Revolving Credit Outstanding Defl by PCE(LFE)	1968:Q1-2014:Q4	5	1	1
170	Real_ConsuCred	Total Consumer Credit Outstanding Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	0
171	FRBSLO_Consumers	FRB Senior Loans Officer Opions. Net Percentage of Domestic Respondents Reporting Increased Willingness to Make Consumer Installment Loans (Fred from 1982:Q2 on Earlier is DB series)	1970:Q1-2014:Q4	1	0	1
	(10) International Variables					
172	Ex rate: major	FRB Nominal Major Currencies Dollar Index (Linked to EXRUS in 1973:1)	1959:Q1-2014:Q4	5	0	1
173	Ex rate: Euro	U.S. / Euro Foreign Exchange Rate	1999:Q1-2014:Q4	5	0	1
174	Ex rate: Switz	Foreign exchange rate: Switzerland (Swiss franc per U.S.$) Fred 1971. EXRSW previous	1971:Q1-2014:Q4	5	0	1
175	Ex rate: Japan	Foreign exchange rate: Japan (yen per U.S.$) Fred 1971- EXRJAN previous	1971:Q1-2014:Q4	5	0	1
176	Ex rate: UK	Foreign exchange rate: United Kingdom (cents per pound) Fred 1971-> EXRUK Previous	1971:Q1-2014:Q4	5	0	1
177	EX rate: Canada	Foreign exchange rate: Canada (Canadian $ per U.S.$) Fred 1971 -> EXRCAN previous	1971:Q1-2014:Q4	5	0	1
178	OECD GDP	OECD: Gross Domestic Product by Expenditure in Constant Prices: Total Gross; Growth Rate (Quartely); Fred Series NAEXKP01O1Q657S	1961:Q2-2013:Q4	1	0	1
179	IP Europe	OECD: Total Ind. Prod (excl Construction) Europe Growth Rate (Quarterly); Fred Series PRINTO01OEQ657S	1960:Q2-2013:Q4	1	0	1
180	Global Ec Activity	Kilian's estimate of glaobal economic activity in industrial commodity markets (Kilian website)	1968:Q1-2014:Q4	1	0	1
	(11) **Asset Prices, Wealth, and Household Balance Sheets**					
181	S&P 500	S&P's Common Stock Price Index: Composite (1941-43=10)	1959:Q1-2014:Q4	5	0	1
182	Real_HHW:TA	Households and nonprofit organizations; total assets (FoF) Seasonally Adjusted (RATS X11) Defl by PCE(LFE)	1959:Q1-2014:Q3	5	0	0

183	Real_HHW:TL	Households and nonprofit organizations; total liabilities Seasonally Adjusted (RATS X11) Defl by PCE(LFE)	1959:Q1-2014:Q3	5	0	1
184	liab_PDI	Liabilities Relative to Person Disp Income	1959:Q1-2014:Q3	5	0	0
185	Real_HHW:W	Households and nonprofit organizations; net worth (FoF) Seasonally Adjusted (RATS X11) Defl by PCE(LFE)	1959:Q1-2014:Q3	5	0	1
186	W_PDI	Networth Relative to Personal Disp Income	1959:Q1-2014:Q3	1	0	0
187	Real_HHW:TFA	Households and nonprofit organizations; total financial assets Seasonally Adjusted (RATS X11) Defl by PCE(LFE)	1959:Q1-2014:Q3	5	0	0
188	Real_HHW:TA_RE	TotalAssets minus Real Estate Assets Defl by PCE(LFE)	1959:Q1-2014:Q3	5	0	1
189	Real_HHW:TNFA	Households and nonprofit organizations; total nonfinancial assets (FoF) Seasonally Adjusted (RATS X11) Defl by PCE(LFE)	1959:Q1-2014:Q3	5	0	0
190	Real_HHW:RE	Households and nonprofit organizations; real estate at market value Seasonally Adjusted (RATS X11) Defl by PCE(LFE)	1959:Q1-2014:Q3	5	0	1
191	DJIA	Common Stock Prices: Dow Jones Industrial Average	1959:Q1-2014:Q4	5	0	1
192	VXO	VXO (Linked by N. Bloom) .. Average daily VIX from 2009 ->	1962:Q3-2014:Q4	1	0	1
193	Real_Hprice:OFHEO	House Price Index for the United States Defl by PCE(LFE)	1975:Q1-2014:Q4	5	0	1
194	Real_CS_10	Case-Shiller 10 City Average Defl by PCE(LFE)	1987:Q1-2014:Q4	5	0	1
195	Real_CS_20	Case-Shiller 20 City Average Defl by PCE(LFE)	2000:Q1-2014:Q4	5	0	1
	(12) Other					
196	Cons. Expectations	Consumer expectations NSA (Copyright University of Michigan)	1959:Q1-2014:Q4	1	0	1
197	PoilcyUncertainty	Baker Bloom Davis Policy Uncertainty Index	1985:Q1-2014:Q4	2	0	1
	(13) Oil Market Variables					
198	World Oil Production	World Oil Production.1994:Q1 on from EIA (Crude Oil including Lease Condensate); Data prior to 1994 from From Baumeister and Peerlman (2013)	1959:Q1-2014:Q3	5	0	0
199	World Oil Production	World Oil Production.1994:Q1 on from EIA (Crude Oil including Lease Condensate); Data prior to 1994 from From Baumeister and Peerlman (2013); Seasonally adjusted using RATS X11 (note seasonality before 1970)	1959:Q1-2014:Q3	5	0	1
200	IP: Energy Prds	IP: Consumer Energy Products	1959:Q1-2014:Q4	5	0	1
201	Petroleum Stocks	U.S. Ending Stocks excluding SPR of Crude Oil and Petroleum Products (Thousand Barrels); SA using X11 in RATS	1959:Q1-2014:Q4	5	0	1
202	Real_Price:Oil	PPI: Crude Petroleum Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	1
203	Real_Crudeoil Price	Crude Oil: West Texas Intermediate (WTI) - Cushing Oklahoma Defl by PCE(LFE)	1986:Q1-2014:Q4	5	0	1
204	Real_CrudeOil	Crude Oil Prices: Brent - Europe Defl by PCE(LFE) Def	1987:Q3-2014:Q4	5	0	1
205	Real_Price Gasoline	Conventional Gasoline Prices: New York Harbor Regular Defl by PCE(LFE)	1986:Q3-2014:Q4	5	0	1
206	Real_Refiners Acq. Cost (Imports)	U.S. Crude Oil Imported Acquisition Cost by Refiners (Dollars per Barrel) Defl by PCE(LFE)	1974:Q1-2014:Q4	5	0	1
207	Real_CPI Gasoline	CPI Gasoline (NSA) BLS: CUUR0000SETB01 Defl by PCE(LFE)	1959:Q1-2014:Q4	5	0	1

Dealing with large datasets

(1) Outliers

(2) Non-stationarities and 'trends'

Usual transformations (logs, differences, spreads, etc.)

Low-frequency 'demeaning'

(3) Aggregates (139 vs. 207)

(4) Estimate factors using standarized data ('weights' in weighted least squares). [$\min_{\{F_t\},\{\lambda_i\}} \sum_{i,t} (X_{it} - \lambda_i' F_t)^2$]

Low-frequency 'demeaning' weights and sprectral gain

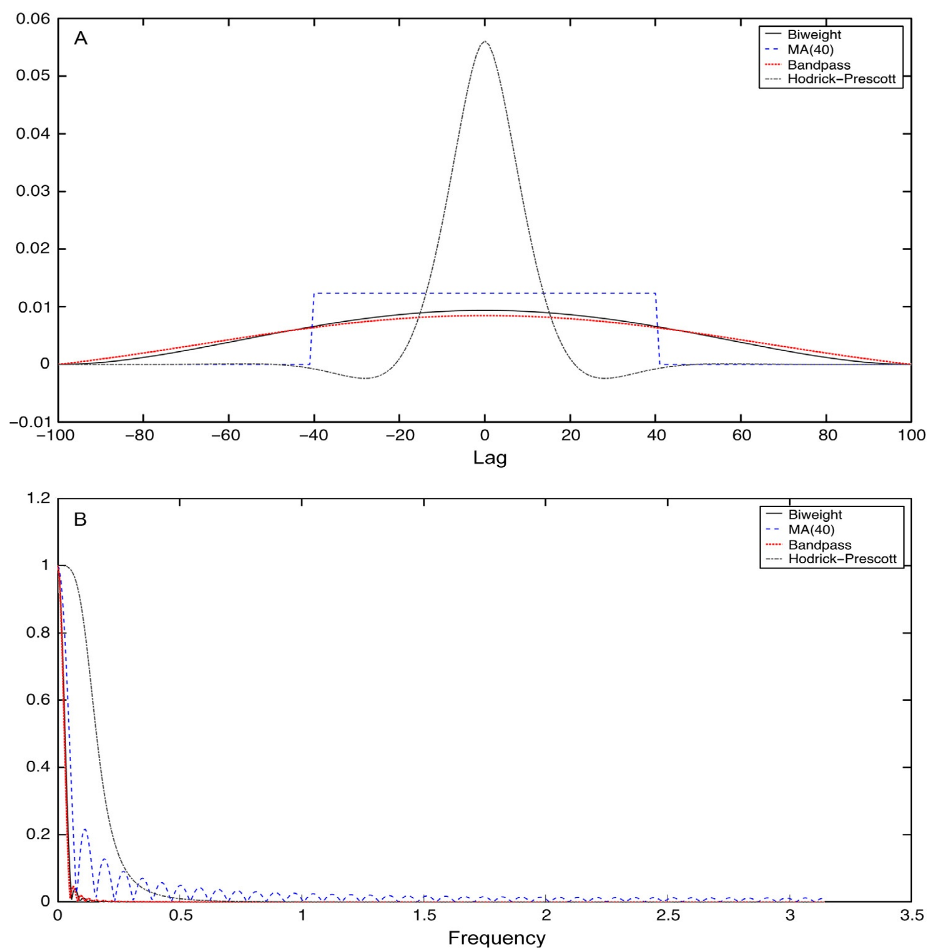


Fig. 2 Lag weights and spectral gain of trend filters. *Notes:* The biweight filter uses a bandwidth (truncation parameter) of 100 quarters. The bandpass filter is a 200-quarter low-pass filter truncated after 100 leads and lags (Baxter and King, 1999). The moving average is equal-weighted with 40 leads and lags. The Hodrick and Prescott (1997) filter uses 1600 as its tuning parameter.

How Many Factors?

(1) Scree plot

(2) Information criteria

(3) Others

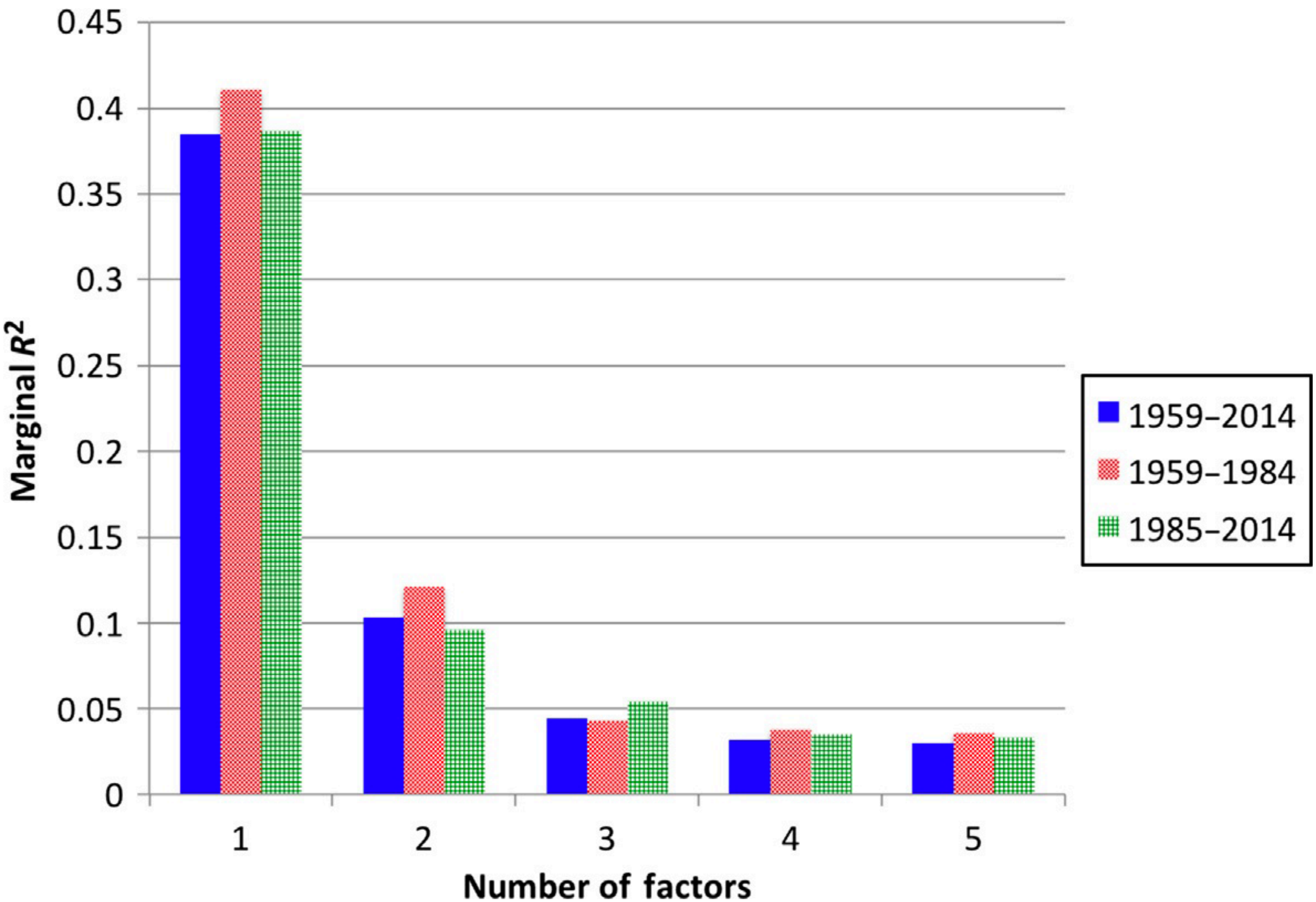
Least squares objective function for r factors:

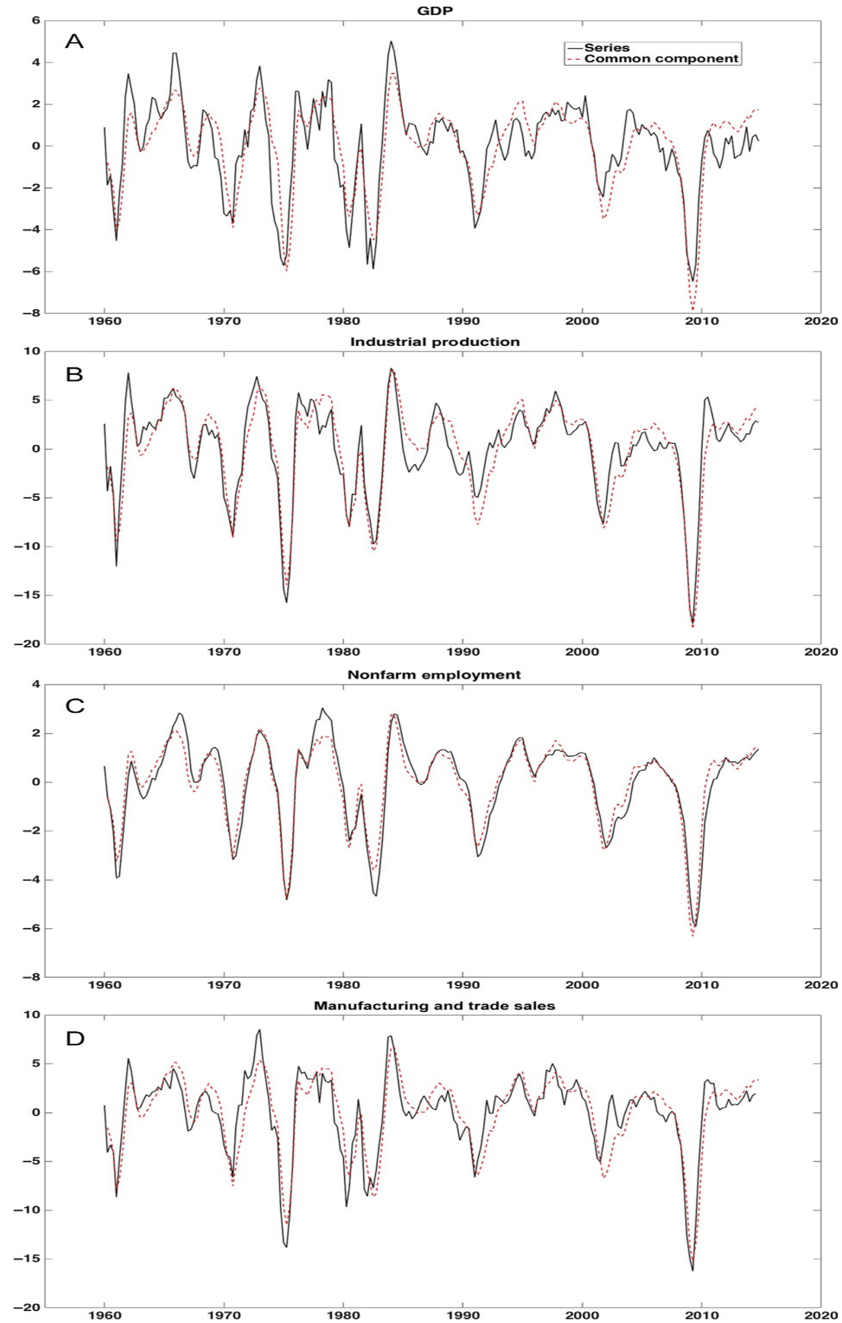
$$SSR(r) = \min_{\{F_t\},\{\lambda_i\}} \sum_{i,t} (X_{it} - \lambda_i ' F_t)^2$$

where F_t and λ_i are $r \times 1$ vectors.

Scree plot: Marginal (trace) R^2 for factor k:

Scree plot for 58 real variables





:trended four-quarter growth rates of US GDP, industrial production, nonfarm

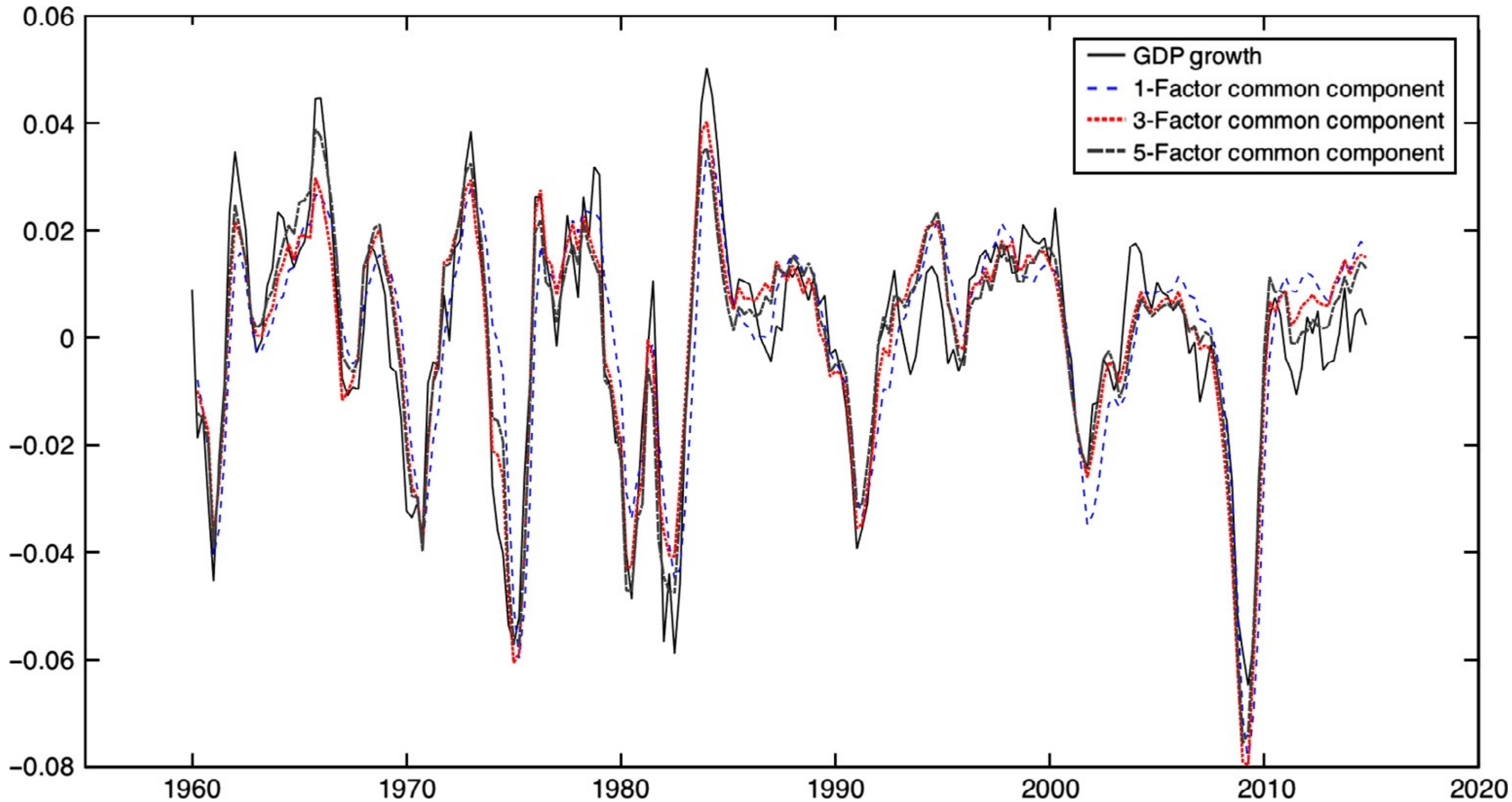
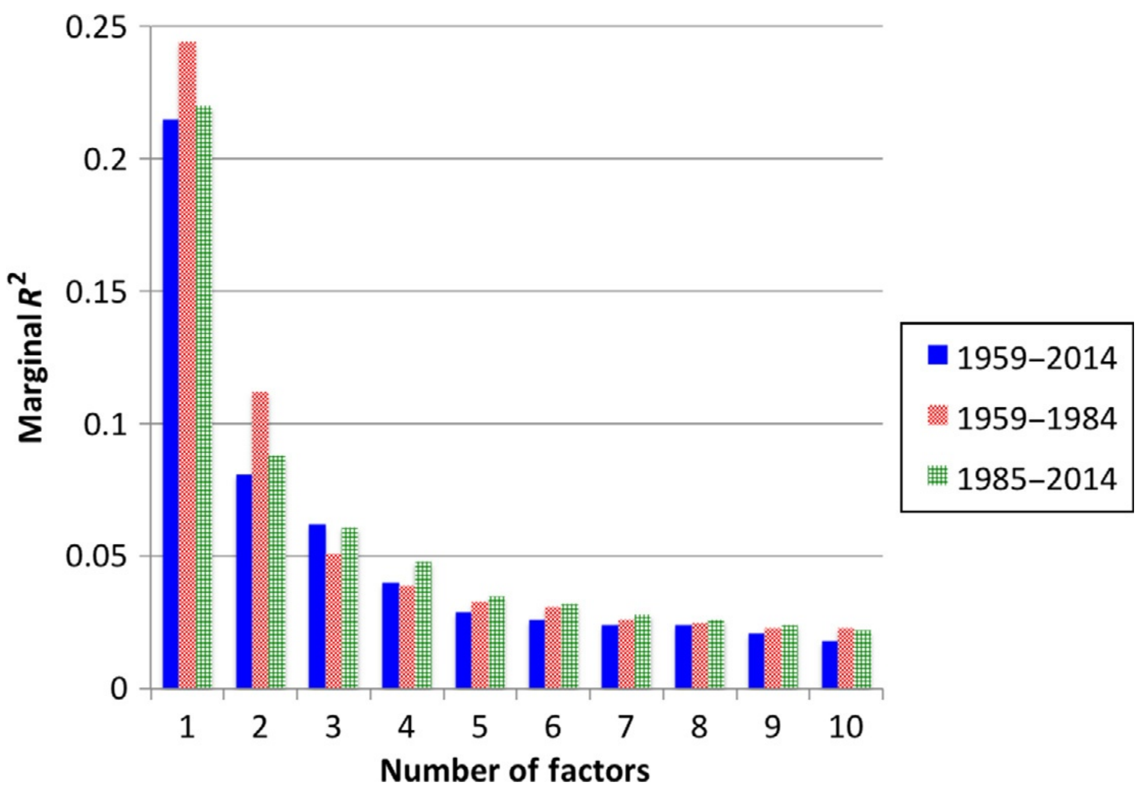


Fig. 4 Four-quarter GDP growth (*black*) and its common component based on 1, 3, and 5 static factors: real activity dataset.

Scree plot – Full data set (139 variables)

Factor Models and Structural Vector Autoregressio



Information criteria: Bai and Ng

$\text{IC}(r) = \ln(\text{SSR}(r)) + rg(\text{sample size})$

Sample size: n and T

$$\text{BNIC}(r) = \ln\left(\text{SSR}(r)\right) + r\left(\frac{n+T}{nT}\right)\ln\left(\min(n,T)\right)$$

Note: when $n = T$ this is $\text{BNIC}(r) = \ln(\text{SSR}(r)) + 2r\times\ln(T)/T$.

Table 2 Statistics for estimating the number of static factors

(A) Real activity dataset ($N = 58$ disaggregates used for estimating factors)

Number of static factors	Trace R^2	Marginal trace R^2	BN-IC_{p2}	AH-ER
1	0.385	0.385	−0.398	**3.739**
2	0.489	0.103	−0.493	2.338
3	0.533	0.044	**−0.494**	1.384
4	0.565	0.032	−0.475	1.059
5	0.595	0.030	−0.458	1.082

(B) Full dataset ($N = 139$ disaggregates used for estimating factors)

Number of static factors	Trace R^2	Marginal trace R^2	BN-IC_{p2}	AH-ER
1	0.215	0.215	−0.183	**2.662**
2	0.296	0.081	−0.233	1.313
3	0.358	0.062	−0.266	1.540
4	0.398	0.040	**−0.271**	1.368
5	0.427	0.029	−0.262	1.127
6	0.453	0.026	−0.249	1.064
7	0.478	0.024	−0.235	1.035
8	0.501	0.024	−0.223	1.151
9	0.522	0.021	−0.205	1.123
10	0.540	0.018	−0.185	1.057

'Static' and 'Dynamic' factors (again)

$$X_t = \lambda(L)f_t + e_t \text{ and } \phi(L)f_t = \eta_t$$

$$X_t = \begin{pmatrix} \lambda_0 & \lambda_1 & \cdots & \lambda_k \end{pmatrix} \begin{pmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-k} \end{pmatrix} + e_t$$

$$\begin{pmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-k} \end{pmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{k+1} \\ 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} f_{t-1} \\ f_{t-2} \\ \vdots \\ f_{t-k-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \eta_t$$

or

$$X_t = \Lambda F_t + e_t$$
$$F_t = \Phi F_{t-1} + G\eta_t$$

Number of static factors (r) = number of elements in F

Number of dynamic factors (q) = number of elements in f = number of elements in η = number of common shocks.

Determining q: Several ways. Here is one:

$$X_t = \Lambda F_t + e_t = \Lambda\eta_t + \beta F_{t\text{-}1} + e_t \quad \text{(with } \beta = \Lambda\Phi\text{)}.$$

$$\Rightarrow$$

Use BNIC on the residuals from the regression of X_t onto $\hat{F}_{t-1}$.

(C) Amenguel-Watson estimate of number of dynamic factors: BN-IC_{pi} values, full dataset ($N = 139$)

No. of dynamic factors	Number of static factors									
	1	2	3	4	5	6	7	8	9	10
1	−0.098	−0.071	−0.072	−0.068	−0.069	−0.065	−0.064	−0.064	−0.064	−0.060
2		**−0.085**	−0.089	−0.087	−0.089	−0.084	−0.084	−0.084	−0.085	−0.080
3			**−0.090**	**−0.088**	**−0.091**	**−0.088**	**−0.088**	**−0.086**	**−0.086**	**−0.084**
4				−0.077	−0.080	−0.075	−0.075	−0.073	−0.072	−0.069
5					−0.064	−0.060	−0.062	−0.057	−0.055	−0.052
6						−0.045	−0.043	−0.040	−0.037	−0.036
7							−0.024	−0.022	−0.020	−0.018
8								−0.002	0.000	0.003
9									0.021	0.023
10										0.044

Notes: BN-IC_{p2} denotes the Bai and Ng (2002) IC_{p2} information criterion. AH-ER denotes the Ahn and Horenstein (2013) ratio of $(i+1)$th to ith eigenvalues. The minimal BN-IC_{p2} entry in each column, and the maximal Ahn–Horenstein ratio entry in each column, is the respective estimate of the number of factors and is shown in bold. In panel C, the BN-IC_{p2} values are computed using the covariance matrix of the residuals from the regression of the variables onto lagged values of the column number of static factors, estimated by principal components.

Table 3 Importance of factors for selected series for various numbers of static and dynamic factors: full dataset DFM

	A. R^2 of common component			B. Fraction of four quarters ahead forecast error variance due to common component		
	Number of static factors r			Number of dynamic factors q with $r = 8$ static factors		
Series	**1**	**4**	**8**	**1**	**4**	**8**
Real GDP	0.54	0.65	0.81	0.39	0.77	0.83
Employment	0.84	0.92	0.93	0.79	0.86	0.90
Housing starts	0.00	0.52	0.67	0.49	0.51	0.75
Inflation (PCE)	0.05	0.51	0.64	0.34	0.66	0.67
Inflation (core PCE)	0.02	0.13	0.17	0.24	0.34	0.41
Labor productivity (NFB)	0.02	0.30	0.59	0.12	0.46	0.54
Real hourly labor compensation (NFB)	0.00	0.25	0.70	0.19	0.67	0.71
Federal funds rate	0.25	0.41	0.54	0.52	0.54	0.62
Ted-spread	0.26	0.59	0.61	0.18	0.33	0.59
Term spread (10 year–3 month)	0.00	0.36	0.72	0.32	0.38	0.63
Exchange rates	0.01	0.22	0.70	0.05	0.60	0.68
Stock prices (SP500)	0.06	0.49	0.73	0.14	0.29	0.79
Real money supply (MZ)	0.00	0.25	0.34	0.15	0.24	0.29
Business loans	0.11	0.49	0.51	0.13	0.16	0.23
Real oil prices	0.04	0.68	0.70	0.40	0.66	0.71
Oil production	0.09	0.10	0.12	0.01	0.04	0.12

(Use VAR(4) and AR(4) for *e*'s to compute forecast error variances)

What about many more factors?
(Full 138-variable dataset)

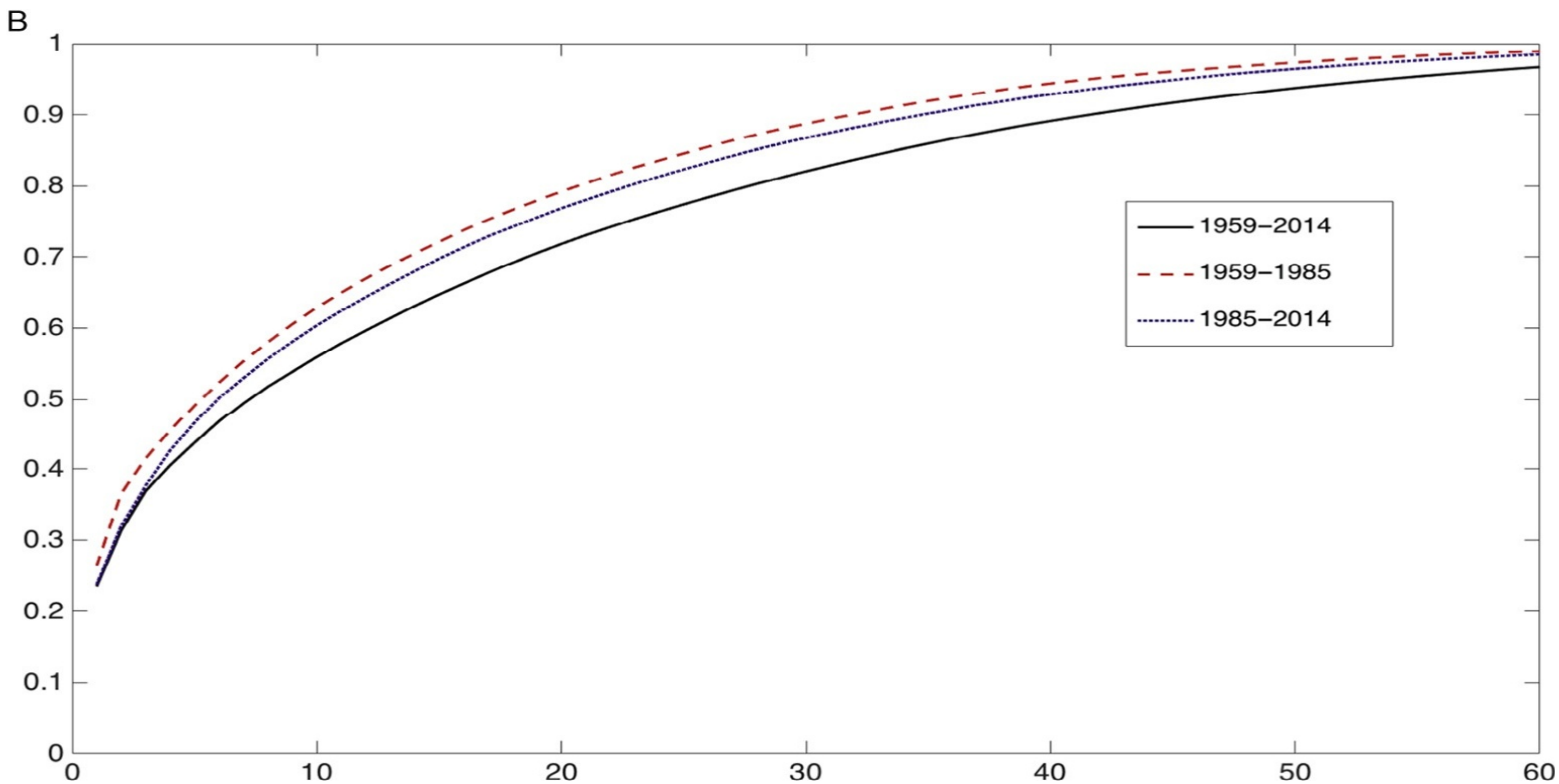


Fig. 6 (A) Scree plot for full dataset: full sample, pre-1984, and post-1984. (B) Cumulative R^2 as a

Is there useful information in additional factors? (For forecasting, maybe)

Instability in Factor Models (references in paper)

Two key results:

(1) Common discrete changes increase the number of factors

(2) Idiosynchratic (or weakly correlated) changes have little effect on estimated factors.

Return to single factor model: $X_{it} = \lambda_{i,t} f_t + e_t$

Result 1:

Suppose $\lambda_{i,t} = \begin{cases} \lambda_{i1} \text{ for } t \le T_1 \\ \lambda_{i2} \text{ for } t > T_1 \end{cases}$ and break is pervasive:

Write

$$X_{it} = (\lambda_{i1} \;\; \lambda_{i2}) \begin{pmatrix} f_{1t} \\ f_{2t} \end{pmatrix} + e_{it} \text{ where}$$

$f_{1t} = \begin{cases} f_t \text{ for } t \le T_1 \\ 0 \text{ for } t > T_1 \end{cases}$ and f_{2t} is defined analogously

$$X_{it} = \lambda_{i,t} f_t + e_t$$

$$\Rightarrow$$

$$\frac{1}{n}\sum_{i=1}^{n} X_{it} = \left(\frac{1}{n}\sum_{i=1}^{n} \lambda_{i,t}\right) f_t + \frac{1}{n}\sum_{i=1}^{n} e_{it}$$

Results 2 follows from this.

Odds and ends:

(1) Testing for breaks in λs. (Chow-tests, sup-Wald (QLR) tests etc.)

(2) Testing for instability of second moments of common components, var(ΛF_t).

(3) What's changing, λ_i or second moments of F_t? (the composite, $\lambda_i F_t$ affects X_{it}). (What changed during Great Recession … Stock-Watson BPEA 2012)

Stability in the 207-variable macro dataset (some results shown already previous figures)

Table 4 Stability tests for the four- and eight-factor full dataset DFMs
(A) Fraction of rejections of stability null hypothesis

Level of test	Chow test (1984q4 break)	QLR test
(i) Four factors		
1%	0.39	0.62
5%	0.54	0.77
10%	0.63	0.83
(ii) Eight factors		
1%	0.55	0.94
5%	0.65	0.98
10%	0.72	0.98

(B) Distribution of correlations between full- and split-sample common components

	Percentile of distribution				
	5%	**25%**	**50%**	**75%**	**5%**
(i) Four factors					
1959–84	0.65	0.89	0.96	0.99	1.00
1985–2014	0.45	0.83	0.95	0.97	0.99
(ii) Eight factors					
1959–84	0.57	0.83	0.92	0.97	0.99
1985–2014	0.43	0.80	0.94	0.97	0.99

(C) Results by category (four factors)

Category	Number of series	Fraction of Chow test rejections for 5% test	Median correlation between full- and split-sample common components	
			1959–84	1985–2014
NIPA	20	0.50	0.98	0.96
Industrial production	10	0.50	0.98	0.97
Employment and unemployment	40	0.40	0.99	0.99
Orders, inventories, and sales	10	0.80	0.98	0.96
Housing starts and permits	8	0.75	0.96	0.91
Prices	35	0.49	0.88	0.90
Productivity and labor earnings	10	0.80	0.92	0.67
Interest rates	12	0.33	0.98	0.94
Money and credit	9	0.89	0.93	0.89
International	3	0.00	0.97	0.97
Asset prices, wealth, and household balance sheets	12	0.58	0.95	0.92
Other	1	1.00	0.95	0.91
Oil market variables	6	0.83	0.79	0.79

Notes: These results are based on the 176 series with data available for at least 80 quarters in both the pre- and post-84 samples. The Chow tests in (A) and (C) test for a break in 1984q4.