Factor models

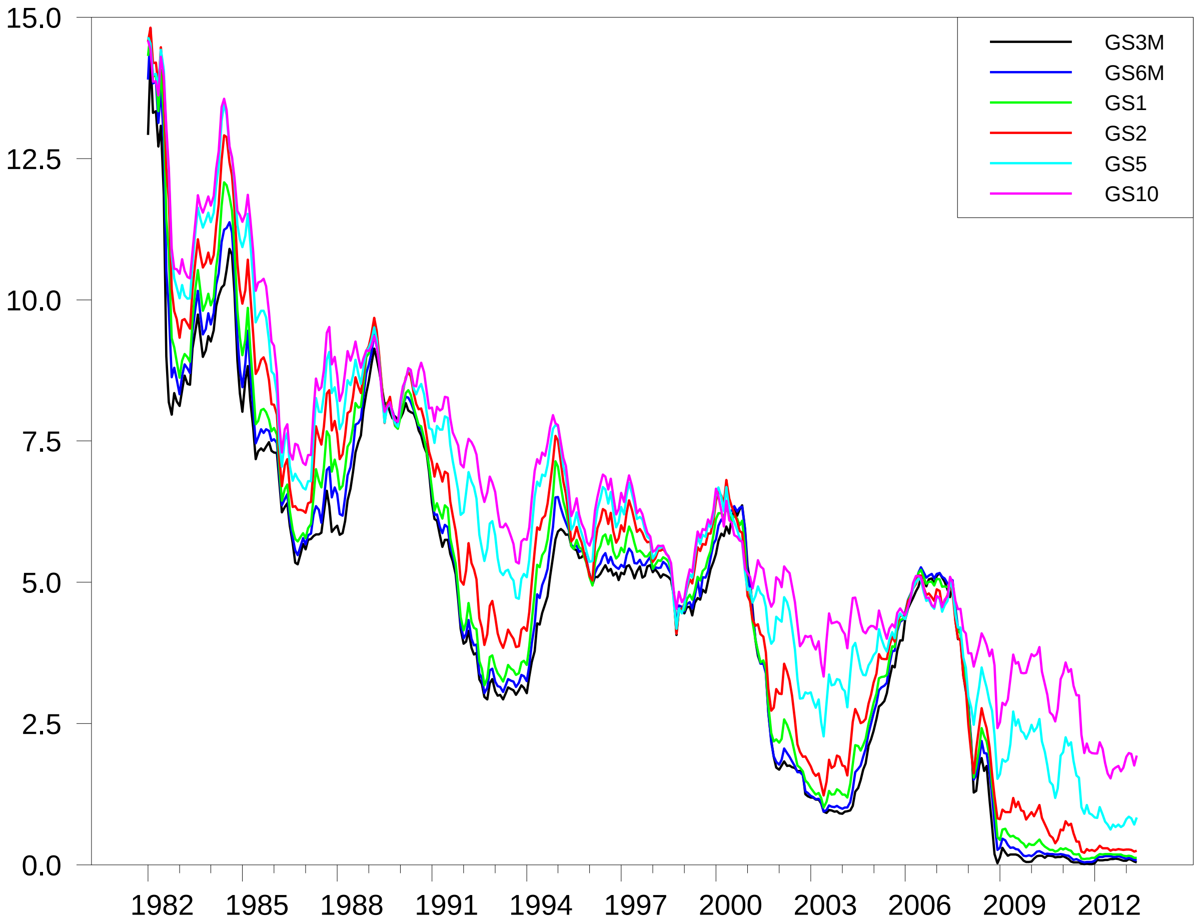
A. Principal components
B. Principal components with missing data
C. Dynamic factor models
D. Factor-augmented vector autoregressions

A. Principal components

Suppose we have a large number of variables observed at date *t*

Goal: can we summarize most of the features of the data using just a few indicators?

Yields on U.S. Treasury securities (3 months to 10 years)



$\mathbf{y}_t = (n \times 1)$ vector of

stationary observations

$\hat{\mu}_i = T^{-1} \sum_{t=1}^{T} y_{it}$ (mean of variable i)

$\hat{\sigma}_{ii} = T^{-1} \sum_{t=1}^{T} (y_{it} - \hat{\mu}_i)^2$

$\tilde{y}_{it} = \hat{\sigma}_{ii}^{-1/2} (y_{it} - \hat{\mu}_i)$

$\tilde{\mathbf{y}}_t = (\tilde{y}_{1t}, \ldots, \tilde{y}_{nt})'$

$\hat{\mathbf{\Omega}} = T^{-1} \sum_{t=1}^{T} \tilde{\mathbf{y}}_t \tilde{\mathbf{y}}_t'$

(sample correlation matrix)

Goal is to find a scalar ξ_t and $(n \times 1)$ vector $\mathbf{h}$ so as to minimize

$$\sum_{t=1}^{T} (\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)'(\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)$$

Note: $\mathbf{h}$ and ξ_t are not unique $(\mathbf{h}\xi_t = \mathbf{h}^* \xi_t^*$ for $\mathbf{h}^* = q\mathbf{h}$, $\xi_t^* = q^{-1}\xi_t)$ but $\mathbf{h}\xi_t$ is unique.

One normalization: $\mathbf{h}'\mathbf{h} = 1$.

$$\tilde{\mathbf{y}}_t \simeq \mathbf{h}\xi_t$$

Scalar ξ_t explains as much of variation of $\tilde{\mathbf{y}}_t$ as possible.

Solution ξ_t^* is called the "first principal component" of $\mathbf{y}_t$ (determined up to arbitrary scale factor).

Elements of vector $\mathbf{h}$ are called "factor loadings".

$$\min_{\{\mathbf{h},\xi_1,\ldots,\xi_T\}} \sum_{t=1}^{T} (\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)'(\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)$$

Concentrate objective function:

(1) for any $\mathbf{h}$, find best $\{\xi_1, \ldots, \xi_T\}$

(2) substitute $\xi_t(\mathbf{h})$ into objective

and min with respect to $\mathbf{h}$

(1) for fixed $\mathbf{h}$:

$$\min_{\{\xi_1,\ldots,\xi_T\}} \sum_{t=1}^{T} (\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)'(\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)$$

$$\min_{\xi_t} \, (\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)'(\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)$$

OLS regression of $\tilde{\mathbf{y}}_t$ on $\mathbf{h}$

$$\xi_t(\mathbf{h}) = (\mathbf{h}'\mathbf{h})^{-1}\mathbf{h}'\tilde{\mathbf{y}}_t$$

$$(\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)'(\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t)$$

$$= \tilde{\mathbf{y}}_t'(\mathbf{I}_n - \mathbf{h}(\mathbf{h}'\mathbf{h})^{-1}\mathbf{h}')\tilde{\mathbf{y}}_t$$

(2) minimize over $\mathbf{h}$:

$$\min_{\{\mathbf{h}\}} \sum_{t=1}^{T} (\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t(\mathbf{h}))'(\tilde{\mathbf{y}}_t - \mathbf{h}\xi_t(\mathbf{h}))$$

$$= \sum_{t=1}^{T} \tilde{\mathbf{y}}_t'(\mathbf{I}_n - \mathbf{h}(\mathbf{h}'\mathbf{h})^{-1}\mathbf{h}')\tilde{\mathbf{y}}_t$$

$$\Leftrightarrow \quad \max_{\{\mathbf{h}\}} \sum_{t=1}^{T} \tilde{\mathbf{y}}_t'\mathbf{h}(\mathbf{h}'\mathbf{h})^{-1}\mathbf{h}'\tilde{\mathbf{y}}_t$$

subject to $\mathbf{h}'\mathbf{h} = 1$

$$= \sum_{t=1}^{T} \mathbf{h}'\tilde{\mathbf{y}}_t\tilde{\mathbf{y}}_t'\mathbf{h}$$

$$= \mathbf{h}'\left(\sum_{t=1}^{T} \tilde{\mathbf{y}}_t\tilde{\mathbf{y}}_t'\right)\mathbf{h}$$

$$= T\mathbf{h}'\hat{\boldsymbol{\Omega}}\mathbf{h}$$

$$\max_{\{\mathbf{h}\}} \mathbf{h}'\hat{\mathbf{\Omega}}\mathbf{h}$$

subject to $\mathbf{h}'\mathbf{h} = 1$

Consider eigenvalues of $\hat{\boldsymbol{\Omega}}$

$$\hat{\boldsymbol{\Omega}}\mathbf{x}_i = \hat{\lambda}_i \mathbf{x}_i \quad \text{for } i = 1,..,n$$

$$\hat{\boldsymbol{\Lambda}} = \text{diag}(\hat{\lambda}_1,..,\hat{\lambda}_n)$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix}$$

$$\mathbf{X}'\mathbf{X} = \mathbf{I}_n$$

$$\hat{\boldsymbol{\Omega}}\mathbf{X} = \mathbf{X}\hat{\boldsymbol{\Lambda}}$$

$$\mathbf{X}'\hat{\boldsymbol{\Omega}}\mathbf{X} = \hat{\boldsymbol{\Lambda}}$$

$$\max_{\{\mathbf{h}\}} \mathbf{h}'\hat{\mathbf{\Omega}}\mathbf{h} \text{ subject to } \mathbf{h}'\mathbf{h} = 1$$

Let $\mathbf{h} = \mathbf{X}\mathbf{h}^*$

where $\mathbf{X}'\mathbf{X} = \mathbf{I}_n$ and $\mathbf{X}'\hat{\mathbf{\Omega}}\mathbf{X} = \hat{\mathbf{\Lambda}}$

$$\max_{\{\mathbf{h}\}} \mathbf{h}'\hat{\mathbf{\Omega}}\mathbf{h} \text{ subject to } \mathbf{h}'\mathbf{h} = 1$$

$$\Leftrightarrow \quad \max_{\{\mathbf{h}^*\}} \mathbf{h}^{*\prime}\mathbf{X}'\hat{\mathbf{\Omega}}\mathbf{X}\mathbf{h}^* \text{ subject to } \mathbf{h}^{*\prime}\mathbf{h}^* = 1$$

$$\mathbf{h}^{*\prime}\mathbf{X}'\hat{\mathbf{\Omega}}\mathbf{X}\mathbf{h}^* = \mathbf{h}^{*\prime}\hat{\mathbf{\Lambda}}\mathbf{h}^*$$

$$= h_1^{*2}\hat{\lambda}_1 + \cdots + h_n^{*2}\hat{\lambda}_n$$

$$\max_{\{\mathbf{h}^*\}} h_1^{*2}\hat{\lambda}_1 + \cdots + h_n^{*2}\hat{\lambda}_n$$

$$\text{s.t. } h_1^{*2} + \cdots + h_n^{*2} = 1$$

$$\text{Solution: } h_1^* = 1$$

$$h_2^* = h_3^* = \cdots = h_n^* = 0$$

$$\Rightarrow \mathbf{h} = \mathbf{x}_1$$

Conclusion: the factor loadings are given by the eigenvector of $\hat{\mathbf{\Omega}}$ associated with largest eigenvalue. The first principal component is given by $\mathbf{h}'\tilde{\mathbf{y}}_t,$ the product of this eigenvector with de-meaned data vector.

Example:

$\mathbf{y}_t$ = interest rates for month t

for U.S. Treasury securities with

maturities 3m, 6m, 1y, 2y, 5y, 10y

$n = 6$

t = 1982:M1 - 2013:M5

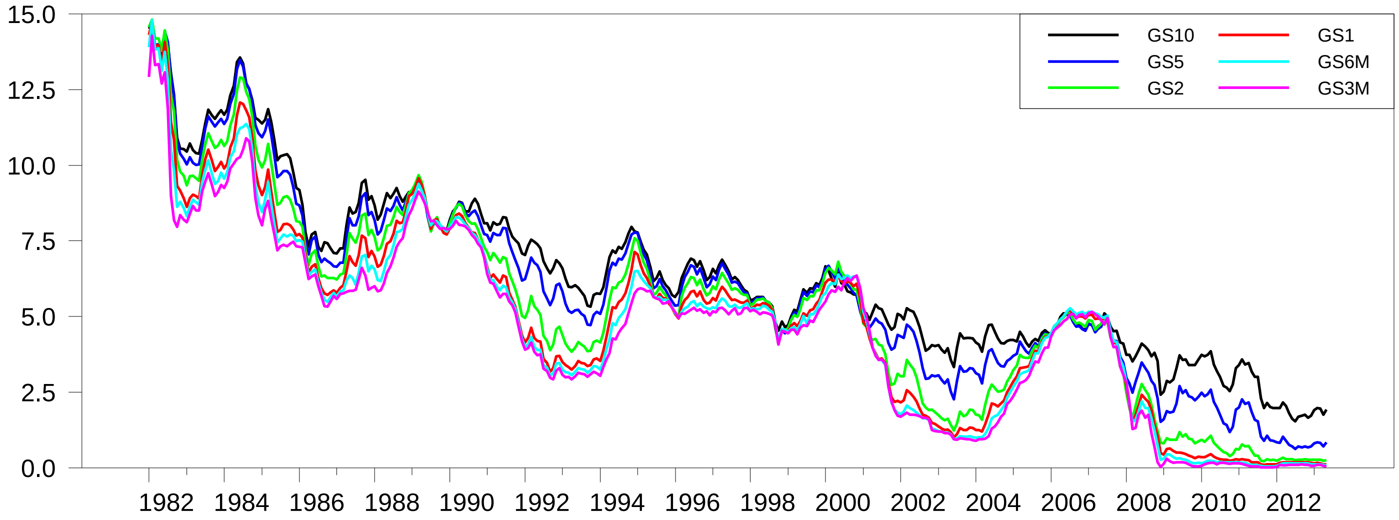
$$\hat{\boldsymbol{\Omega}} = T^{-1} \sum_{t=1}^{T} (\mathbf{y}_t - \hat{\boldsymbol{\mu}})(\mathbf{y}_t - \hat{\boldsymbol{\mu}})'$$

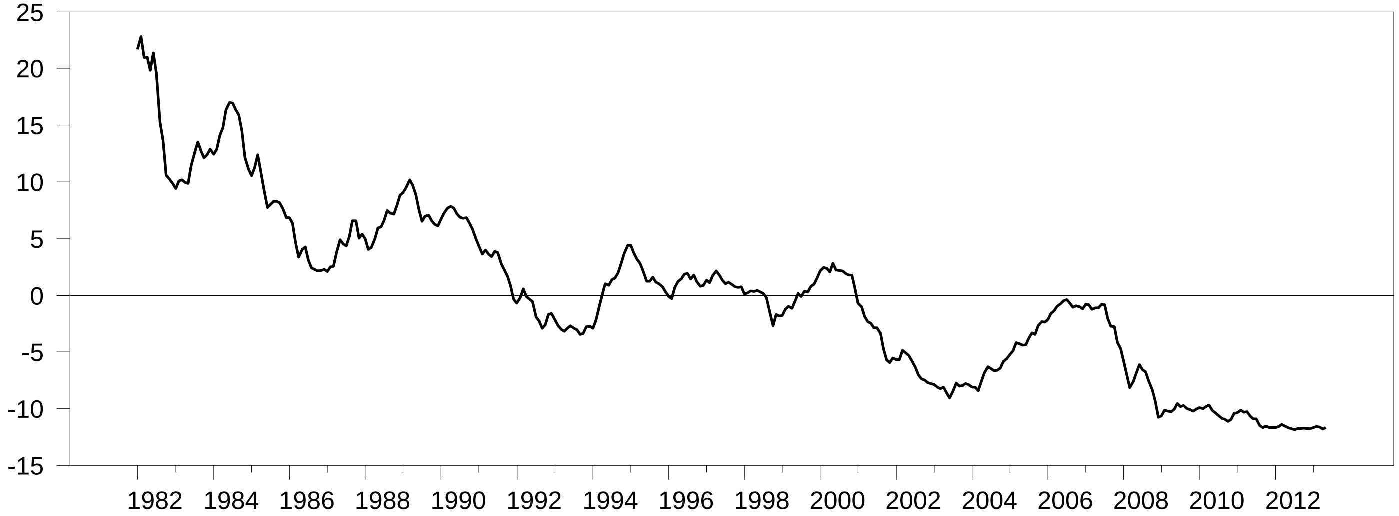
Eigenvector of $\hat{\boldsymbol{\Omega}}$ associated with largest eigenvalue:

(0.3999, 0.4153, 0.4244, 0.4344, 0.4061, 0.3659)$'$

$1/\sqrt{6} = 0.4082$

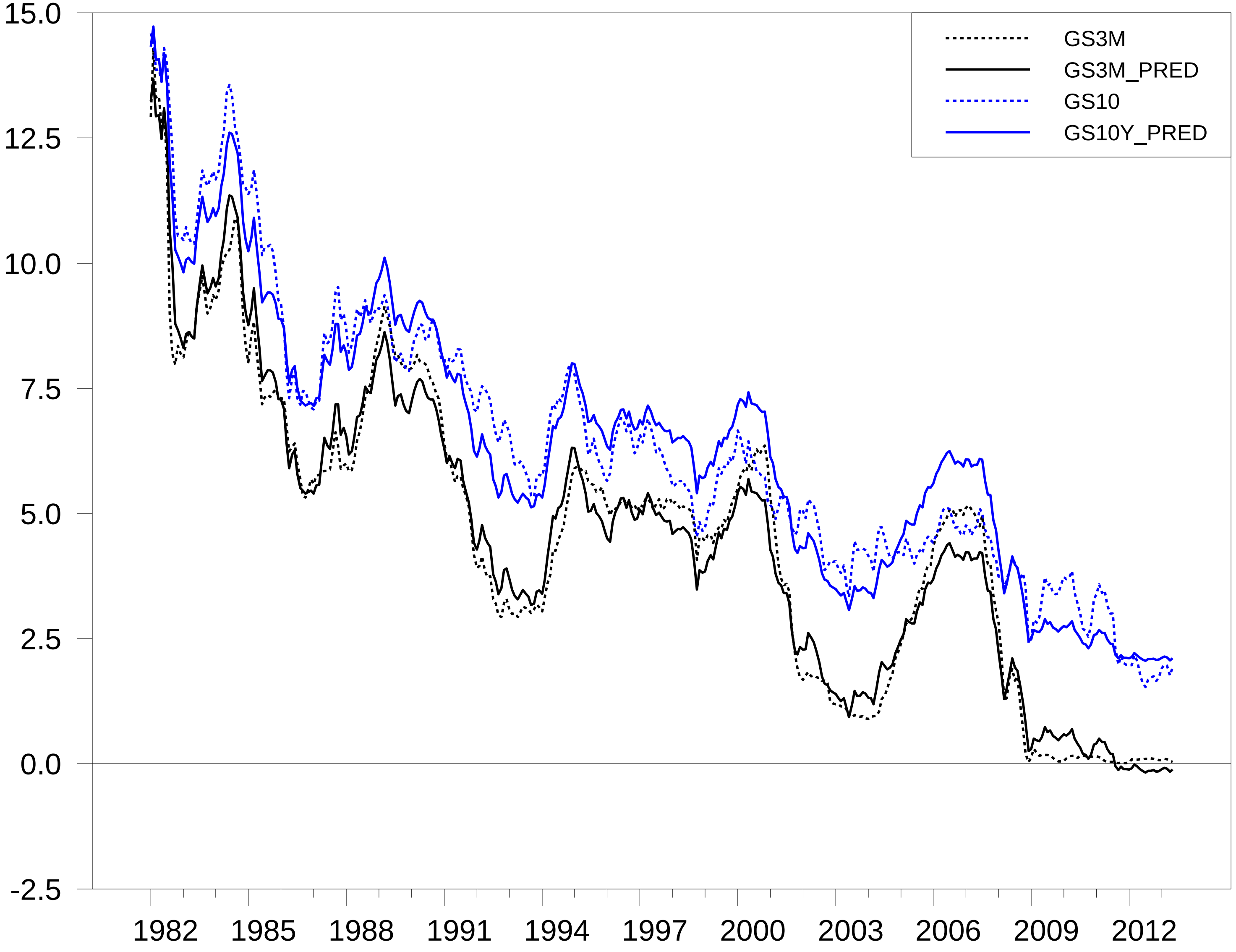
Conclusion: first principal component is essentially the average of the 6 yields.


15.0
12.5
10.0
7.5
5.0
2.5
0.0
GS10
GS5
GS2
GS1
GS6M
GS3M
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012


25
20
15
10
5
0
-5
-10
-15
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012

Fitted value for yield i:

$$y_{it} \simeq \hat{\mu}_i + h_i \hat{\xi}_t$$


15.0
12.5
10.0
7.5
5.0
2.5
0.0
-2.5
GS3M
GS3M_PRED
GS10
GS10Y_PRED
1982
1985
1988
1991
1994
1997
2000
2003
2006
2009
2012

Could also ask: suppose I could use 2 variables to summarize the 6 yields. Choose (2×1) vector $\boldsymbol{\xi}_t$ for $t = 1, \ldots, T$ and $(n \times 2)$ matrix $\mathbf{H}$ to minimize $\sum_{t=1}^{T} (\tilde{\mathbf{y}}_t - \mathbf{H}\boldsymbol{\xi}_t)'(\tilde{\mathbf{y}}_t - \mathbf{H}\boldsymbol{\xi}_t)$.

Again not unique:

$\mathbf{Q}$ nonsingular (2×2) matrix

$$\mathbf{H}^* = \mathbf{HQ}$$

$$\boldsymbol{\xi}_t^* = \mathbf{Q}^{-1}\boldsymbol{\xi}_t$$

$$\mathbf{H}\boldsymbol{\xi}_t = \mathbf{H}^*\boldsymbol{\xi}_t^*.$$

Normalize $\mathbf{H}'\mathbf{H} = \mathbf{I}_2$.

$$\min_{\{\boldsymbol{\xi}_t\}} (\tilde{\mathbf{y}}_t - \mathbf{H}\boldsymbol{\xi}_t)'(\tilde{\mathbf{y}}_t - \mathbf{H}\boldsymbol{\xi}_t)$$

$$= \tilde{\mathbf{y}}_t'(\mathbf{I}_n - \mathbf{H}(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}')\tilde{\mathbf{y}}_t$$

$$\max_{\{\mathbf{H}\}} \sum_{t=1}^{T} \tilde{\mathbf{y}}_t'\mathbf{H}(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'\tilde{\mathbf{y}}_t$$

$$= \sum_{t=1}^{T} \text{trace}\left[(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'\tilde{\mathbf{y}}_t\tilde{\mathbf{y}}_t'\mathbf{H}\right]$$

$$= T\text{trace}\left[\mathbf{H}'\hat{\boldsymbol{\Omega}}\mathbf{H}\right]$$

Solution: $\mathbf{H}$ is in the linear space spanned by the eigenvectors of $\hat{\mathbf{\Omega}}$ associated with the two largest eigenvalues.

Second principal component refers to $\mathbf{h}_2'\tilde{\mathbf{y}}_t$ for $\mathbf{h}_2$ the eigenvector of $\hat{\mathbf{\Omega}}$ associated with the second largest eigenvalue. Note second PC is orthogonal to the first:

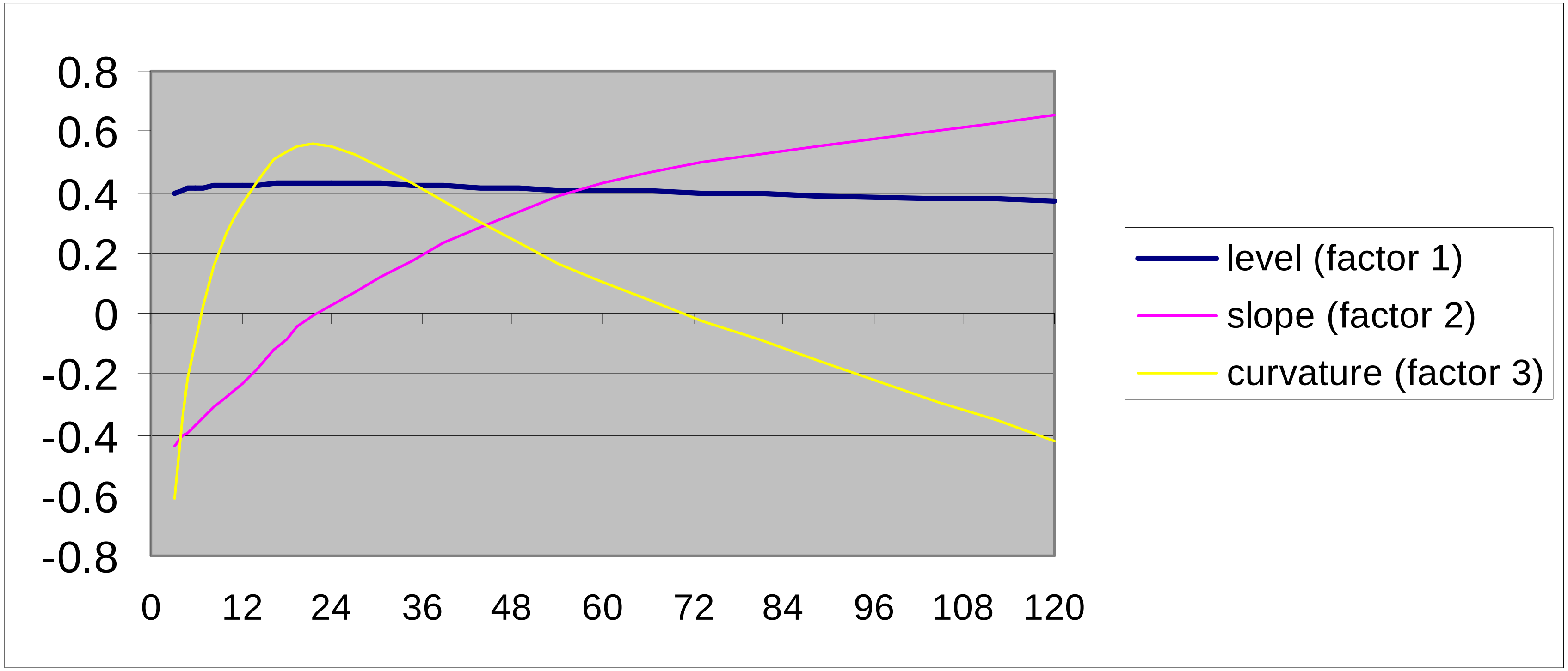
$$\sum_{t=1}^{T}(\mathbf{h}_1'\tilde{\mathbf{y}}_t)(\tilde{\mathbf{y}}_t'\mathbf{h}_2) = T\mathbf{h}_1'\hat{\mathbf{\Omega}}\mathbf{h}_2$$

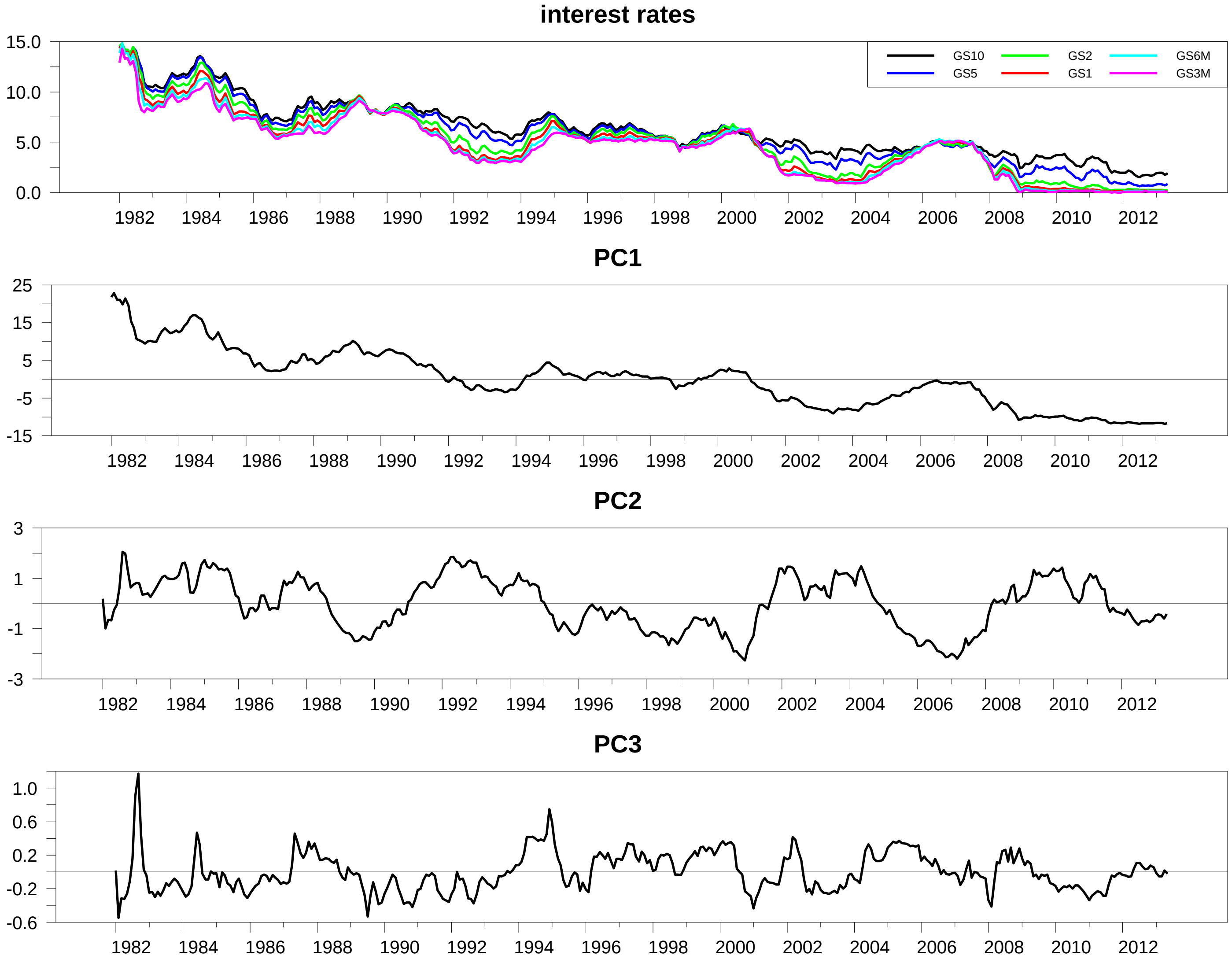
$$= T\hat{\lambda}_2\mathbf{h}_1'\mathbf{h}_2 = 0$$

Interest rates: eigenvalues of $\boldsymbol{\Omega}$

	Eigenvalue	Percent
1	56.0309	0.980548
2	1.0423	0.998789
3	0.0519	0.999697
4	0.013	0.999924
5	3.05E-03	0.999978
6	1.28E-03	1

Factor loadings associated with first three principal components




interest rates
GS10
GS5
GS2
GS1
GS6M
GS3M
PC1
PC2
PC3

Selecting the number of factors r

for standard principal components:

$$V_r = \sum_{t=1}^{T} \Big[\tilde{\mathbf{y}}_t - \underset{(n\times r)}{\mathbf{H}^{(r)*}} \underset{(r\times 1)}{\boldsymbol{\xi}_t^{(r)*}} \Big]' \Big[\tilde{\mathbf{y}}_t - \underset{(n\times r)}{\mathbf{H}^{(r)*}} \underset{(r\times 1)}{\boldsymbol{\xi}_t^{(r)*}} \Big]$$

$$\{\underset{(n\times r)}{\mathbf{H}^{(r)*}}, \underset{(r\times 1)}{\boldsymbol{\xi}_t^{(r)*}}\} =$$

$$\arg \min_{\{\mathbf{H},\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_T\}} (\tilde{\mathbf{y}}_t - \mathbf{H}\boldsymbol{\xi}_t)'(\tilde{\mathbf{y}}_t - \mathbf{H}\boldsymbol{\xi}_t)$$

subject to $\mathbf{H}'\mathbf{H} = \mathbf{I}_r$

Bai and Ng, Econometrica (2002):

Choose r to minimize

$$\log V_r + r\frac{(n+T)\log[\min(n,T)]}{nT}$$

Ahn and Horenstein,

Econometrica (2013):

$$\hat{\boldsymbol{\Omega}} = T^{-1} \sum_{t=1}^{T} \tilde{\mathbf{y}}_t \tilde{\mathbf{y}}_t'$$

$\hat{\lambda}_1$ = largest eigenvalue of $\hat{\boldsymbol{\Omega}}$

$\vdots$

$\hat{\lambda}_n$ = smallest eigenvalue of $\hat{\boldsymbol{\Omega}}$

Choose r to be value for which

$\hat{\lambda}_r / \hat{\lambda}_{r+1}$ is largest.

B. Principal components with missing data

May not observe all n variables at date t

Want to summarize using $\underset{(r\times 1)}{\boldsymbol{\xi}_t}$

$s_{it} = 1$ if variable i observed at t, 0 if not

$\underset{(1\times r)}{\mathbf{h}'_{i,.}} =$ row i of $\underset{(n\times r)}{\mathbf{H}}$

$$\min_{\{\mathbf{h}_{1,.},\ldots,\mathbf{h}_{n,.},\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_T\}} \sum_{t=1}^{T} \sum_{i=1}^{n} (\tilde{y}_{it} - \mathbf{h}'_{i,.}\boldsymbol{\xi}_t)^2 s_{it}$$

Stock-Watson iterative algorithm

(1) Given an estimate of $\mathbf{H}$, estimate $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_T$ by T separate cross-section OLS regressions of $\tilde{y}_{it}$ on $\mathbf{h}_{i,.}$

Regression t minimizes $\sum_{i=1}^{n} (\tilde{y}_{it} - \mathbf{h}_{i,.}' \boldsymbol{\xi}_t)^2 s_{it}$

$$\Rightarrow \hat{\boldsymbol{\xi}}_t = \left(\sum_{i=1}^{n} \mathbf{h}_{i,.} \mathbf{h}_{i,.}' s_{it}\right)^{-1} \left(\sum_{i=1}^{n} \mathbf{h}_{i,.} y_{it} s_{it}\right)$$

(2) Given an estimate of $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_T\}$, estimate $\mathbf{h}_{\cdot,i}$ by n separate time-series OLS regressions of $\tilde{y}_{it}$ on $\boldsymbol{\xi}_t$

Regression i minimizes

$$\sum_{t=1}^{T} (\tilde{y}_{it} - \mathbf{h}'_{i,\cdot}\boldsymbol{\xi}_t)^2 s_{it}$$

$$\Rightarrow \hat{\mathbf{h}}_{i,\cdot} = \left(\sum_{t=1}^{T} \boldsymbol{\xi}_t \boldsymbol{\xi}'_t s_{it}\right)^{-1} \left(\sum_{t=1}^{T} \boldsymbol{\xi}_t \tilde{y}_{it} s_{it}\right)$$

(3) Iterate on (1) and (2) starting iteration from $\mathbf{H}$ based on eigenvectors of correlation matrix for subset of data for which all observations available.

(4) Normalize final estimates by finding

$\underset{(k\times k)}{\mathbf{Q}}$ matrix of eigenvectors of $T^{-1}\sum_{t=1}^{T}\hat{\boldsymbol{\xi}}_t\hat{\boldsymbol{\xi}}_t'$

$$\boldsymbol{\xi}_t^* = \mathbf{Q}'\hat{\boldsymbol{\xi}}_t$$

$$\Rightarrow T^{-1}\sum_{t=1}^{T}\boldsymbol{\xi}_t^*\boldsymbol{\xi}_t^{*\prime} = \mathbf{Q}'T^{-1}\sum_{t=1}^{T}\hat{\boldsymbol{\xi}}_t\hat{\boldsymbol{\xi}}_t'\mathbf{Q}$$

$$= \mathbf{Q}'\mathbf{Q}\hat{\boldsymbol{\Lambda}} = \hat{\boldsymbol{\Lambda}} \text{ (diagonal)}$$

Factor models

A. Principal components
B. Principal components with missing data
C. Dynamic factor models

$\mathbf{y}_t$ = observed variables

$\boldsymbol{\xi}_t$ = unobserved factors

$$\underset{(n\times 1)}{\mathbf{y}_t} = \underset{(n\times r)}{\mathbf{H}} \underset{(r\times 1)}{\boldsymbol{\xi}_t} + \underset{(n\times 1)}{\mathbf{u}_t}$$

$$\underset{(r\times 1)}{\boldsymbol{\xi}_{t+1}} = \underset{(r\times r)}{\boldsymbol{\Phi}} \underset{(r\times 1)}{\boldsymbol{\xi}_t} + \underset{(r\times 1)}{\mathbf{v}_t}$$

$$\underset{(n\times 1)}{\mathbf{u}_{t+1}} = \underset{(n\times n)}{\mathbf{D}} \underset{(n\times 1)}{\mathbf{u}_t} + \underset{(n\times 1)}{\boldsymbol{\varepsilon}_t}$$

$\mathbf{D}$ = diag$(d_1, d_2, \ldots, d_n)$

$$\underset{(n\times 1)}{\mathbf{y}_t} = \underset{(n\times r)}{\mathbf{H}} \underset{(r\times 1)}{\boldsymbol{\xi}_t} + \underset{(n\times 1)}{\mathbf{u}_t}$$

Assumption 1:

$$n^{-1}\mathbf{H}'\mathbf{H} \underset{n\to\infty}{\to} \underset{(r\times r)}{\mathbf{Q}_{\mathbf{H}}}$$

with rank($\mathbf{Q}_{\mathbf{H}}$) $= r$.

Means factors matter for more

than just finite subset of $\mathbf{y}_t$

and are different from each other

(columns of $\mathbf{H}$ not too similar).

Assumption 2:

maximum eigenvalue of $E(\mathbf{u}_t\mathbf{u}_t')$

is $\leq c$ for all n.

Means $\mathbf{u}_t$ does not have its own

factor structure.

(e.g., for $E(\mathbf{u}_t\mathbf{u}_t') = \sigma^2\mathbf{1}\mathbf{1}'$

then $\mathbf{1}$ is eigenvector with

eigenvalue $\sigma^2\mathbf{1}'\mathbf{1} = \sigma^2 n$)

Suppose these assumptions held and there was an $(n \times r)$ matrix $\mathbf{W}$ such that:

(i) $n^{-1}\mathbf{W}'\mathbf{W} \to \mathbf{I}_r$

(ii) $n^{-1}\mathbf{W}'\mathbf{H} \to \underset{(r \times r)}{\mathbf{\Lambda}}$

(iii) rank($\mathbf{\Lambda}$) $= r$

For yields example and $r = 1$,

$$\mathbf{W}' = (1, 1, \ldots, 1)$$

$$\underset{(n\times 1)}{\mathbf{y}_t} = \underset{(n\times r)}{\mathbf{H}} \underset{(r\times 1)}{\boldsymbol{\xi}_t} + \underset{(n\times 1)}{\mathbf{u}_t}$$

$$n^{-1} \underset{(r\times n)}{\mathbf{W}'} \underset{(n\times 1)}{\mathbf{y}_t} = n^{-1} \underset{(r\times n)}{\mathbf{W}'} \underset{(n\times r)}{\mathbf{H}} \underset{(r\times 1)}{\boldsymbol{\xi}_t}$$

$$+ \; n^{-1} \underset{(r\times n)}{\mathbf{W}'} \underset{(n\times 1)}{\mathbf{u}_t}$$

$$n^{-1}\mathbf{W}'\mathbf{u}_t \overset{p}{\to} \mathbf{0}$$

$$\text{(e.g., } n^{-1}\sum_{i=1}^{n} u_{it} \overset{p}{\to} 0\text{)}$$

$$n^{-1}\mathbf{W}'\mathbf{y}_t = n^{-1}\mathbf{W}'\mathbf{H}\boldsymbol{\xi}_t + n^{-1}\mathbf{W}'\mathbf{u}_t$$

$$n^{-1}\mathbf{W}'\mathbf{y}_t \xrightarrow{p} \boldsymbol{\Lambda}\boldsymbol{\xi}_t$$

Conclusion: can uncover space spanned by $\boldsymbol{\xi}_t$ from $n^{-1}\mathbf{W}'\mathbf{y}_t$.

Don’t need to use any knowledge of dynamics to uncover the variable that drives all the dynamics.

Stock and Watson (JASA, 2002) showed that under related assumptions, the first r principal components of $\mathbf{y}_t$ provide a consistent estimate of $\mathbf{\Lambda\xi}_t$ for some nonsingular $(r \times r)$ matrix $\mathbf{\Lambda}$.

D. Factor-augmented vector autoregressions (FAVAR-- Bernanke, Boivin and Eliasz, QJE, 2005)

$\mathbf{y}_t = (n \times 1)$ vector of observed variables $(n = 120)$

$\mathbf{x}_t = (m \times 1)$ subset of $\mathbf{y}_t$ of special interest or importance.

BBE take $\mathbf{x}_t = r_t$ (the fed funds rate) or $\mathbf{x}_t =$ fed funds rate, industrial production, and inflation, in deviations from their means.

Factor-Augmented VAR:

$$\begin{bmatrix} \underset{(r\times 1)}{\boldsymbol{\xi}_t} \\ \underset{(m\times 1)}{\mathbf{x}_t} \end{bmatrix} = \begin{bmatrix} \underset{(r\times r)}{\boldsymbol{\Phi}_{11}(L)} & \underset{(r\times m)}{\boldsymbol{\Phi}_{12}(L)} \\ \underset{(m\times r)}{\boldsymbol{\Phi}_{21}(L)} & \underset{(m\times m)}{\boldsymbol{\Phi}_{22}(L)} \end{bmatrix} \begin{bmatrix} \underset{(r\times 1)}{\boldsymbol{\xi}_t} \\ \underset{(m\times 1)}{\mathbf{x}_t} \end{bmatrix}$$

$$+ \begin{bmatrix} \underset{(r\times 1)}{\boldsymbol{\varepsilon}_{1t}} \\ \underset{(m\times 1)}{\boldsymbol{\varepsilon}_{2t}} \end{bmatrix}$$

$$\boldsymbol{\Phi}_{ij}(L) = \boldsymbol{\Phi}_{ij}^{(1)} L^1 + \boldsymbol{\Phi}_{ij}^{(2)} L^2 + \cdots + \boldsymbol{\Phi}_{ij}^{(p)} L^p$$

Could estimate space spanned by $\boldsymbol{\xi}_t$ by that spanned by $\hat{\boldsymbol{\xi}}_t$, the first r principal components of $\mathbf{y}_t$.

Question: how to identify monetary policy shock?

Note: since r_t is included in $\mathbf{y}_t$, each element of $\hat{\boldsymbol{\xi}}_t = \mathbf{H}\mathbf{y}_t$ is linear function of r_t.

Claim: a monetary policy shock does not affect "slow-moving variables" (wages, prices) in the current month.

$\mathbf{y}_t^* =$ subset of $\mathbf{y}_t$ that is "slow-moving"

$\hat{\boldsymbol{\xi}}_t =$ first r PC of $\mathbf{y}_t$

$\hat{\boldsymbol{\xi}}_t^* =$ first r PC of $\mathbf{y}_t^*$

(1) regress $\hat{\xi}_{it} = \boldsymbol{\beta}_i' \hat{\boldsymbol{\xi}}_t^* + \alpha_i r_t + e_{it}$

for $i = 1, \ldots, r$

(2) Calculate $\tilde{\xi}_{it} = \hat{\xi}_{it} - \hat{\alpha}_i r_t$

(3) Estimate VAR for $\tilde{\mathbf{x}}_t = (\tilde{\boldsymbol{\xi}}_t', r_t)'$

$\tilde{\mathbf{x}}_t = \boldsymbol{\Phi}(L)\tilde{\mathbf{x}}_t + \boldsymbol{\varepsilon}_t$

(4) Calculate nonorthogonalized impulse-response function

$$\mathbf{\Psi}(L) = [\mathbf{I}_{r+1} - \mathbf{\Phi}(L)]^{-1}$$

$$\mathbf{\Psi}_s = \frac{\partial \tilde{\mathbf{x}}_{t+s}}{\partial \boldsymbol{\varepsilon}_t'}$$

and Cholesky factorization

$$\hat{\mathbf{\Omega}} = T^{-1} \sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t' = \hat{\mathbf{P}} \hat{\mathbf{P}}'$$

(5) Effect of monetary policy shock (u_t^M) on $\tilde{\mathbf{x}}_{t+x}$ is

$$\frac{\partial \tilde{\mathbf{x}}_{t+s}}{\partial u_t^M} = \hat{\mathbf{\Psi}}_s \hat{\mathbf{p}}_{r+1}$$

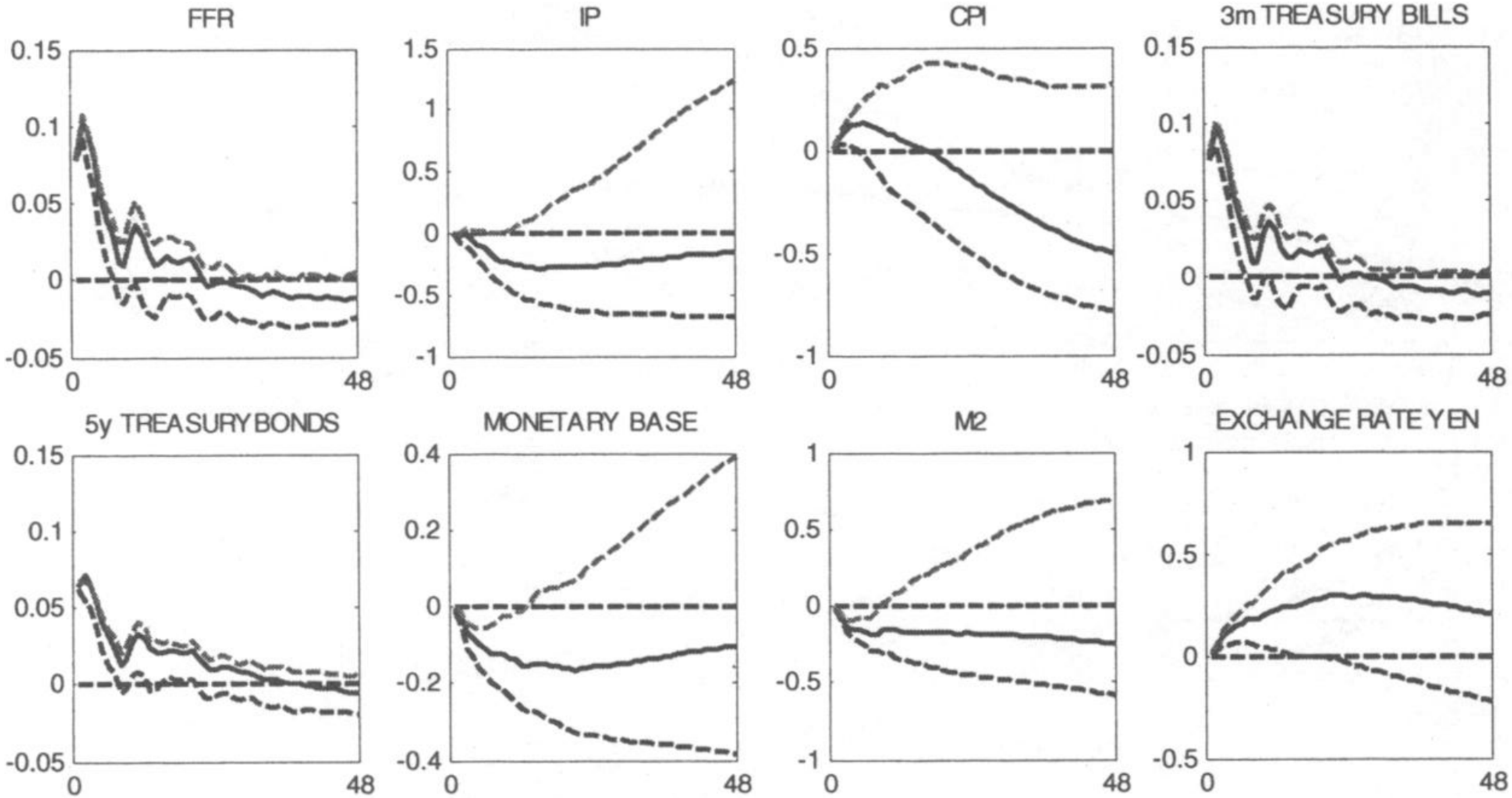
for $\hat{\mathbf{p}}_{r+1}$ the last column of $\hat{\mathbf{P}}$

(6) Since $\underset{(n\times 1)}{\mathbf{y}_t} = \underset{(n\times r)}{\mathbf{H}_\xi} \underset{(r\times 1)}{\tilde{\boldsymbol{\xi}}_t} + \underset{(n\times 1)}{\mathbf{h}_r} r_t$,

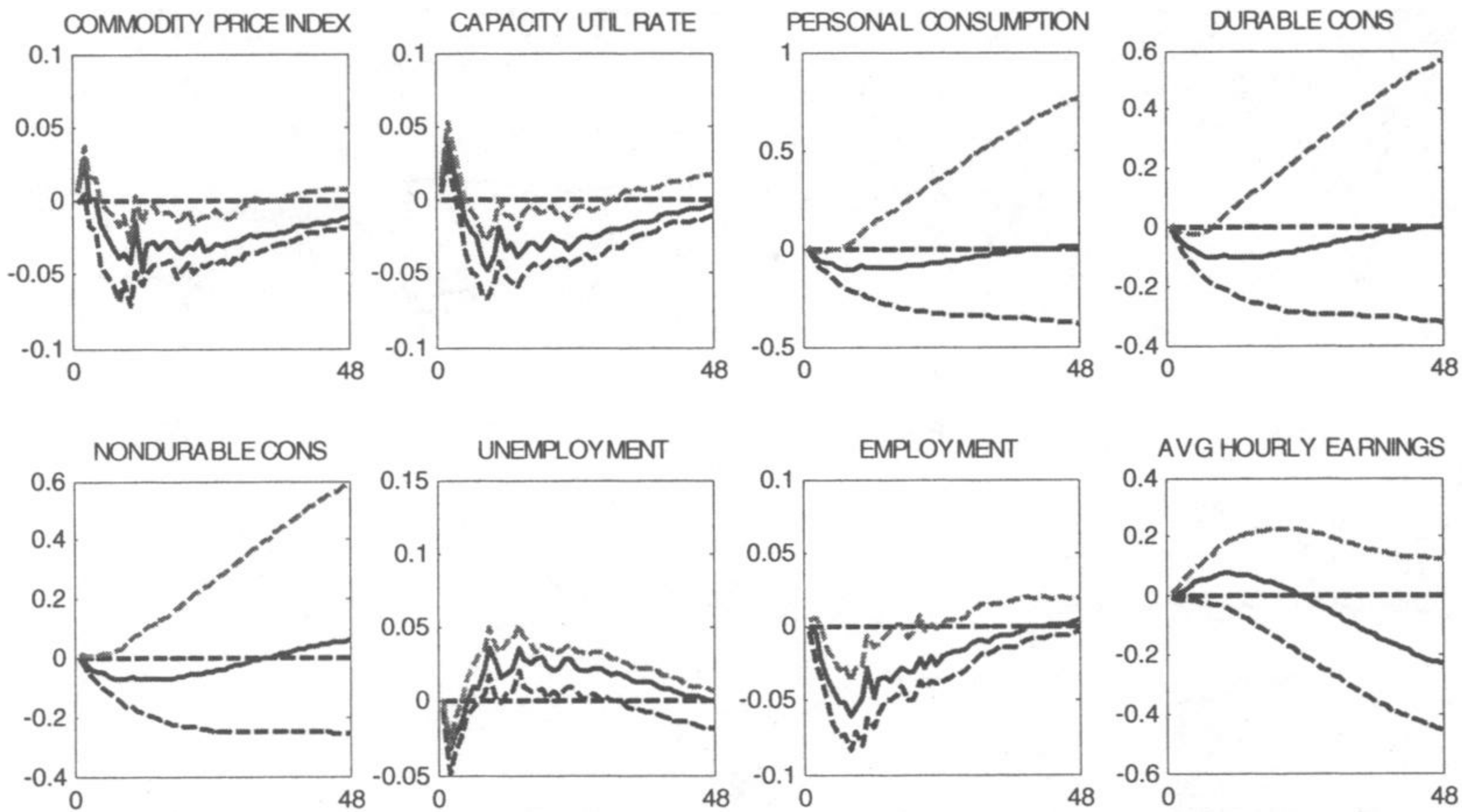
effect of monetary policy on any variable is

$$\frac{\partial \mathbf{y}_{t+s}}{\partial u_t^M} = \begin{bmatrix} \mathbf{H}_\xi & \mathbf{h}_r \end{bmatrix} \hat{\boldsymbol{\Psi}}_s \hat{\mathbf{p}}_{r+1}$$

Effects of monetary policy shock



Effects of monetary policy shock



Effects of monetary policy shock