Priors for the long run

Domenico Giannone
New York Fed

Michele Lenza
European Central Bank

Giorgio Primiceri
Northwestern University

19th Annual DNB Research Conference

September 30, 2016

What we do

- Propose a class of prior distributions for VARs that discipline the long-run implications of the model

↓

Priors for the long run

What we do

- Propose a class of prior distributions for VARs that discipline the long-run implications of the model

↓

Priors for the long run

- Properties
  - Based on macroeconomic theory
  - Conjugate → Easy to implement and combine with existing priors

- Perform well in applications
  - Good (long-run) forecasting performance

Outline

- A specific pathology of (flat-prior) VARs
  - Too much explanatory power of initial conditions and deterministic trends
  - Sims (1996 and 2000)

- Priors for the long run
  - Intuition
  - Specification and implementation

- Alternative interpretations and relation with the literature

- Application: macroeconomic forecasting

Simple example

- AR(1): $$y_t = c + \rho y_{t-1} + \varepsilon_t$$

Simple example

- AR(1): $$y_t = c + \rho y_{t-1} + \varepsilon_t$$
- Iterate backwards: $$y_t = \rho^t y_0 + \sum_{j=0}^{t-1} \rho^j c \ + \ \sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j}$$

Simple example

- AR(1): $$y_t = c + \rho y_{t-1} + \varepsilon_t$$
- Iterate backwards: $$y_t = \underbrace{\rho^t y_0 + \sum_{j=0}^{t-1} \rho^j c}_{\text{DC}} + \underbrace{\sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j}}_{\text{SC}}$$

- Model separates observed variation of the data into
  - DC: deterministic component, predictable from data at time 0
  - SC: unpredictable/stochastic component

Simple example

- AR(1): $$y_t = c + \rho y_{t-1} + \varepsilon_t$$

- Iterate backwards: $$y_t = \underbrace{\rho^t y_0 + \sum_{j=0}^{t-1} \rho^j c}_{\text{DC}} + \underbrace{\sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j}}_{\text{SC}}$$

- Model separates observed variation of the data into
  - DC: deterministic component, predictable from data at time 0
  - SC: unpredictable/stochastic component

- If ***ρ = 1***, DC is a simple linear trend: $$DC = y_0 + c \cdot t$$

Simple example

- AR(1): $$y_t = c + \rho y_{t-1} + \varepsilon_t$$

- Iterate backwards: $$y_t = \underbrace{\rho^t y_0 + \sum_{j=0}^{t-1} \rho^j c}_{\text{DC}} + \underbrace{\sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j}}_{\text{SC}}$$

- Model separates observed variation of the data into
  - DC: deterministic component, predictable from data at time 0
  - SC: unpredictable/stochastic component

- If ***ρ = 1***, DC is a simple linear trend: $$DC = y_0 + c \cdot t$$

- Otherwise more complex: $$DC = \frac{c}{1-\rho} + \rho^t \left(y_0 - \frac{c}{1-\rho} \right)$$

Pathology of (flat-prior) VARs (Sims, 1996 and 2000)

- OLS/MLE has a tendency to “use” the complexity of deterministic components to fit the low frequency variation in the data
- Possible because inference is typically conditional on $\boldsymbol{y_0}$
  - No penalization for parameter estimates of implying steady states or trends far away from initial conditions

Deterministic components in VARs

- Problem more severe with VARs
  - implied deterministic component is much more complex than in AR(1) case

Deterministic components in VARs

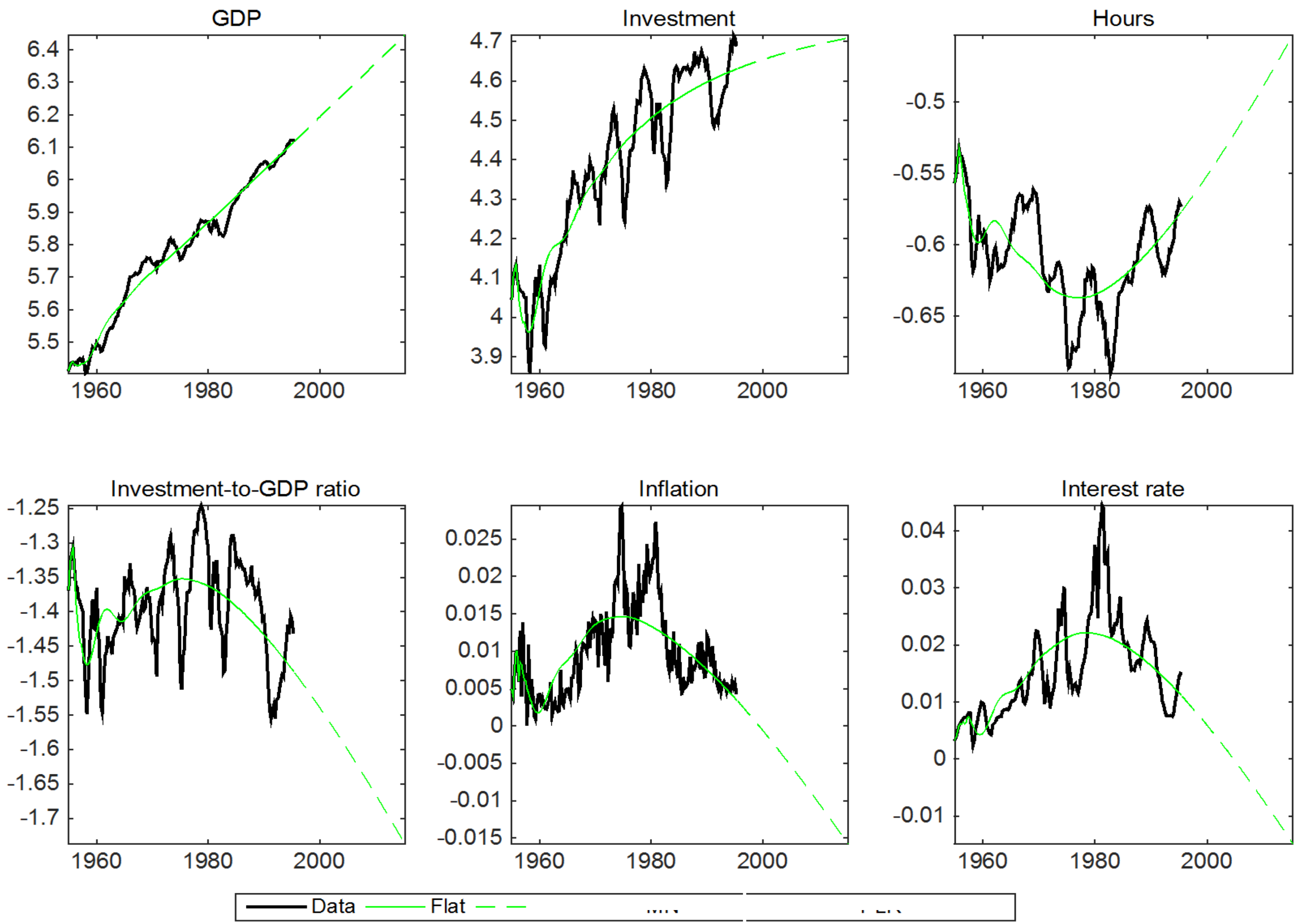
- Problem more severe with VARs
  - implied deterministic component is much more complex than in AR(1) case

- Example: 7-variable VAR(5) with quarterly data on
  - GDP
  - Consumption
  - Investment
  - Real Wages
  - Hours
  - Inflation
  - Federal funds rate

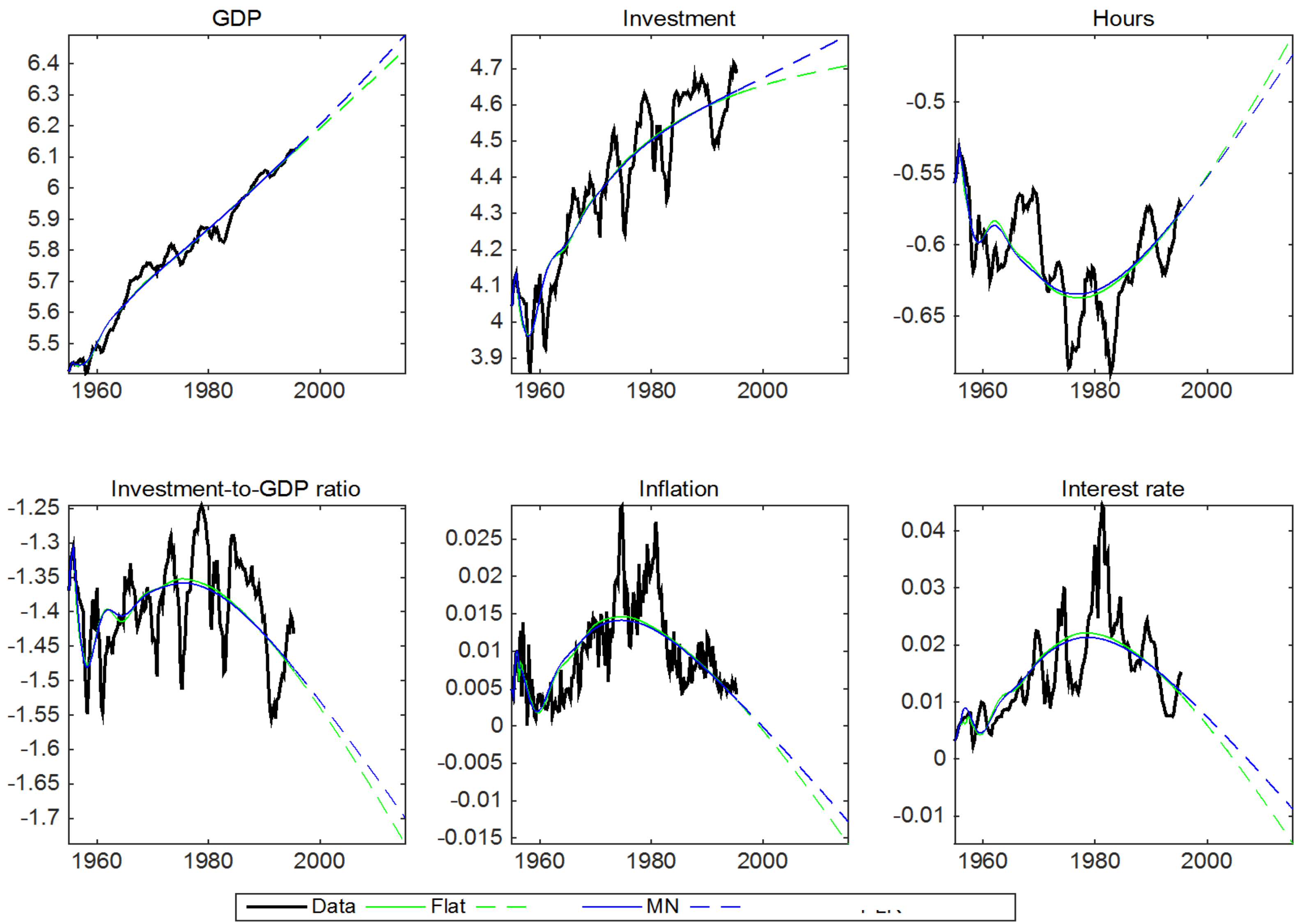
- Sample: 1955:I – 1994:IV

- Flat or Minnesota prior

"Over-fitting" of deterministic components in VARs



“Over-fitting” of deterministic components in VARs



Pathology of (flat-prior) VARs (Sims, 1996 and 2000)

- OLS/MLE has a tendency to “use” the complexity of deterministic components to fit the low frequency variation in the data

- Possible because inference is typically conditional on $\boldsymbol{y_0}$
  - No penalization for parameter estimates of implying steady states or trends far away from initial conditions

➠ Flat-prior VARs attribute an (implausibly) large share of the low frequency variation in the data to deterministic components

Pathology of (flat-prior) VARs (Sims, 1996 and 2000)

- OLS/MLE has a tendency to "use" the complexity of deterministic components to fit the low frequency variation in the data

- Possible because inference is typically conditional on $\boldsymbol{y_0}$
  - No penalization for parameter estimates of implying steady states or trends far away from initial conditions

➠ Flat-prior VARs attribute an (implausibly) large share of the low frequency variation in the data to deterministic components

⇓

- Need a prior that downplays excessive explanatory power of initial conditions and deterministic component

- One solution: center prior on "non-stationarity"

Outline

- A specific pathology of (flat-prior) VARs
  - Too much explanatory power of initial conditions and deterministic trends
  - Sims (1996 and 2000)

- Priors for the long run
  - Intuition
  - Specification and implementation

- Alternative interpretations and relation with the literature

- Application: macroeconomic forecasting

Prior for the long run

$$VAR(1): \quad y_t = c + By_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0,\Sigma)$$

Prior for the long run

$$VAR(1): \quad y_t = c + By_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0,\Sigma)$$

- Rewrite the VAR in terms of levels and differences:

$$\Delta y_t = c + \Pi y_{t-1} + \varepsilon_t$$

$$\Pi = B - I$$

Prior for the long run

$$VAR(1): \quad y_t = c + By_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0,\Sigma)$$

- Rewrite the VAR in terms of levels and differences:

$$\Delta y_t = c + \Pi y_{t-1} + \varepsilon_t$$
$$\Pi = B - I$$

- Prior for the long run ⇒ prior on Π centered at 0

Prior for the long run

$$VAR(1): \quad y_t = c + By_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0,\Sigma)$$

- Rewrite the VAR in terms of levels and differences:

$$\Delta y_t = c + \Pi y_{t-1} + \varepsilon_t$$
$$\Pi = B - I$$

- Prior for the long run ⇒ prior on Π centered at 0

- Standard approach (DLS, SZ, and many followers)
  - Push coefficients towards all variables being independent random walks

Prior for the long run

$$\Delta y_t = c + \Pi y_{t-1} + \varepsilon_t$$

- Rewrite as

$$\Delta y_t = c + \underbrace{\Pi \ H^{-1}}_{\Lambda} \ \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

Prior for the long run

$$\Delta y_t = c + \Pi y_{t-1} + \varepsilon_t$$

- Rewrite as

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- Choose ***H*** and put prior on ***Λ*** conditional on ***H***

Prior for the long run

$$\Delta y_t = c + \Pi y_{t-1} + \varepsilon_t$$

- Rewrite as

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- Choose ***H*** and put prior on ***Λ*** conditional on ***H***

- Economic theory suggests that some linear combinations of ***y*** are less(more) likely to exhibit long-run trends

Prior for the long run

$$\Delta y_t = c + \Pi y_{t-1} + \varepsilon_t$$

- Rewrite as

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- Choose ***H*** and put prior on ***Λ*** conditional on ***H***

- Economic theory suggests that some linear combinations of ***y*** are less(more) likely to exhibit long-run trends

- Loadings associated with these combinations are less(more) likely to be 0

Example: 3-variable VAR of KPSW

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

$$\tilde{y}_{t-1} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{pmatrix} \text{Output} \\ \text{Consumption} \\ \text{Investment} \end{pmatrix}$$

Example: 3-variable VAR of KPSW

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

$$\tilde{y}_{t-1} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{pmatrix} \text{Output} \\ \text{Consumption} \\ \text{Investment} \end{pmatrix}$$

$$\begin{bmatrix} \Delta x_t \\ \Delta c_t \\ \Delta i_t \end{bmatrix} = c + \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} + c_{t-1} + i_{t-1} \\ c_{t-1} - x_{t-1} \\ i_{t-1} - x_{t-1} \end{bmatrix} + \varepsilon_t$$

Example: 3-variable VAR of KPSW

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

$$\tilde{y}_{t-1} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{pmatrix} \text{Output} \\ \text{Consumption} \\ \text{Investment} \end{pmatrix}$$

$$\begin{bmatrix} \Delta x_t \\ \Delta c_t \\ \Delta i_t \end{bmatrix} = c + \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} + c_{t-1} + i_{t-1} \\ c_{t-1} - x_{t-1} \\ i_{t-1} - x_{t-1} \end{bmatrix} + \varepsilon_t$$

Possibly stationary linear combinations

Example: 3-variable VAR of KPSW

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

$$\tilde{y}_{t-1} = \overbrace{\begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{pmatrix} \text{Output} \\ \text{Consumption} \\ \text{Investment} \end{pmatrix}}$$

Common trend

$$\begin{bmatrix} \Delta x_t \\ \Delta c_t \\ \Delta i_t \end{bmatrix} = c + \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} + c_{t-1} + i_{t-1} \\ c_{t-1} - x_{t-1} \\ i_{t-1} - x_{t-1} \end{bmatrix} + \varepsilon_t$$

Possibly stationary linear combinations

Example: 3-variable VAR of KPSW

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

$$\tilde{y}_{t-1} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{pmatrix} \text{Output} \\ \text{Consumption} \\ \text{Investment} \end{pmatrix}$$

Common trend

$$\begin{bmatrix} \Delta x_t \\ \Delta c_t \\ \Delta i_t \end{bmatrix} = c + \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} + c_{t-1} + i_{t-1} \\ c_{t-1} - x_{t-1} \\ i_{t-1} - x_{t-1} \end{bmatrix} + \varepsilon_t$$

Possibly stationary linear combinations

Prior for the long run: specification and implementation

$$\Delta y_t = c + \underbrace{\Pi \ H^{-1}}_{\Lambda} \ \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- $\Lambda_{\cdot i} \mid H, \Sigma \sim N\left(0, \quad \frac{\Sigma}{\left(H_{i\cdot} y_0\right)^2} \right), \qquad i = 1, \ldots, n$

Prior for the long run: specification and implementation

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- $\Lambda_{\cdot i} \,|\, H, \Sigma \sim N\left(0, \phi_i^2 \dfrac{\Sigma}{\left(H_{i\cdot} y_0\right)^2} \right), \qquad i = 1, \ldots, n$

Prior for the long run: specification and implementation

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- $\Lambda_{\cdot i} \mid H, \Sigma \sim N\left(0, \phi_i^2 \frac{\Sigma}{\left(H_{i\cdot} y_0\right)^2} \right), \qquad i = 1, \ldots, n$

- Conjugate
  - Can implement it with Theil mixed estimation in the VAR in levels

Prior for the long run: specification and implementation

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- $\Lambda_{\cdot i} \mid H, \Sigma \sim N\left(0, \phi_i^2 \frac{\Sigma}{\left(H_{i\cdot} y_0\right)^2} \right), \qquad i = 1, \ldots, n$

- Conjugate
  - Can implement it with Theil mixed estimation in the VAR in levels
  - Can be easily combined with existing priors

Prior for the long run: specification and implementation

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- $\Lambda_{\cdot i} \mid H, \Sigma \sim N\left(0, \phi_i^2 \frac{\Sigma}{\left(H_{i\cdot} y_0\right)^2} \right), \qquad i = 1, \ldots, n$

- Conjugate
  - Can implement it with Theil mixed estimation in the VAR in levels
  - Can be easily combined with existing priors
  - Can compute the ML in closed form
    - Useful for hierarchical modeling and setting of hyperparameters ϕ (GLP, 2013)

Empirical results

- Deterministic component in 7-variable VAR

- Forecasting
  - 3-variable VAR
  - 5-variable VAR
  - 7-variable VAR

Empirical results

- Deterministic component in 7-variable VAR

Empirical results

- Deterministic component in 7-variable VAR
  - GDP, Consumption, Investment, Real Wages, Hours, Inflation, Interest Rate

Empirical results

- Deterministic component in 7-variable VAR
  - GDP, Consumption, Investment, Real Wages, Hours, Inflation, Interest Rate

- **H =**

$$
\begin{bmatrix}
Y & C & I & W & H & \pi & R \\
1 & 1 & 1 & 1 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1
\end{bmatrix}
\begin{matrix}
\\
\longrightarrow \\
\longrightarrow \\
\longrightarrow \\
\longrightarrow \\
\longrightarrow \\
\longrightarrow \\
\longrightarrow
\end{matrix}
$$

<u>Interpretation of ***H y***</u>

- Real trend
- Consumption-to-GDP ratio
- Investment-to-GDP ratio
- Labor share
- Hours
- Real interest rate
- Nominal trend

Empirical results

- Deterministic component in 7-variable VAR
  - GDP, Consumption, Investment, Real Wages, Hours, Inflation, Interest Rate

- Forecasting
  - **3-variable VAR**
  - 5-variable VAR
  - 7-variable VAR

- **H =**

$$\begin{bmatrix} Y & C & I & W & H & \pi & R \\ \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{matrix} \\ \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \end{matrix}$$

Interpretation of ***H y***

- Real trend
- Consumption-to-GDP ratio
- Investment-to-GDP ratio
- Labor share
- Hours
- Real interest rate
- Nominal trend

Empirical results

- Deterministic component in 7-variable VAR
  - GDP, Consumption, Investment, Real Wages, Hours, Inflation, Interest Rate

- Forecasting
  - 3-variable VAR
  - **5-variable VAR**
  - 7-variable VAR

- **H =**

$$
H = \begin{bmatrix} Y & C & I & W & H & \pi & R \\ \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}
\begin{matrix} \\ \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \end{matrix}
$$

<u>Interpretation of ***H y***</u>

Real trend
Consumption-to-GDP ratio
Investment-to-GDP ratio
Labor share
Hours
Real interest rate
Nominal trend

Empirical results

- Deterministic component in 7-variable VAR
  - GDP, Consumption, Investment, Real Wages, Hours, Inflation, Interest Rate

- Forecasting
  - 3-variable VAR
  - 5-variable VAR
  - **7-variable VAR**

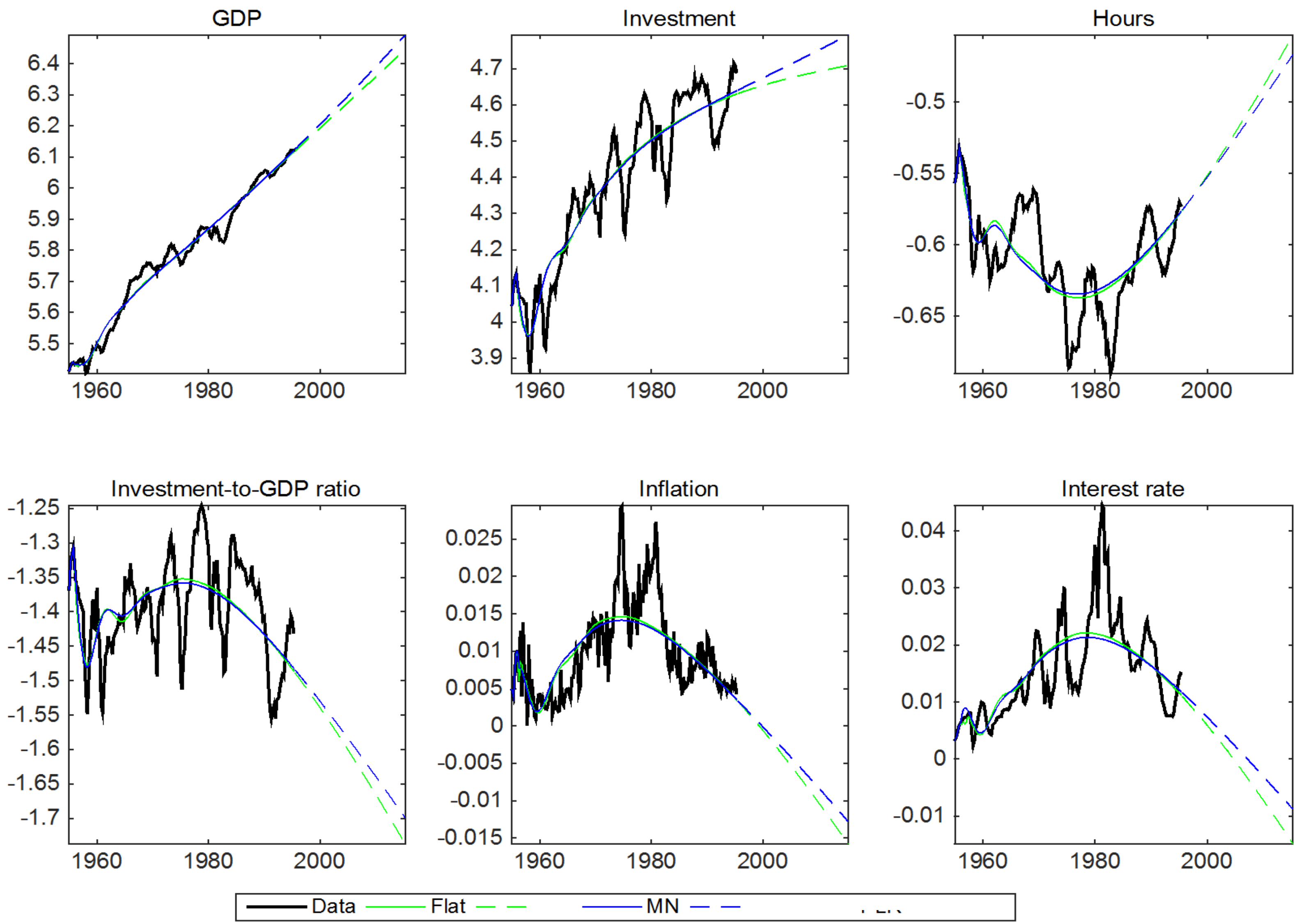
- **H =**

$$\begin{bmatrix} Y & C & I & W & H & \pi & R \\ \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{matrix} \\ \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \\ \longrightarrow \end{matrix}$$

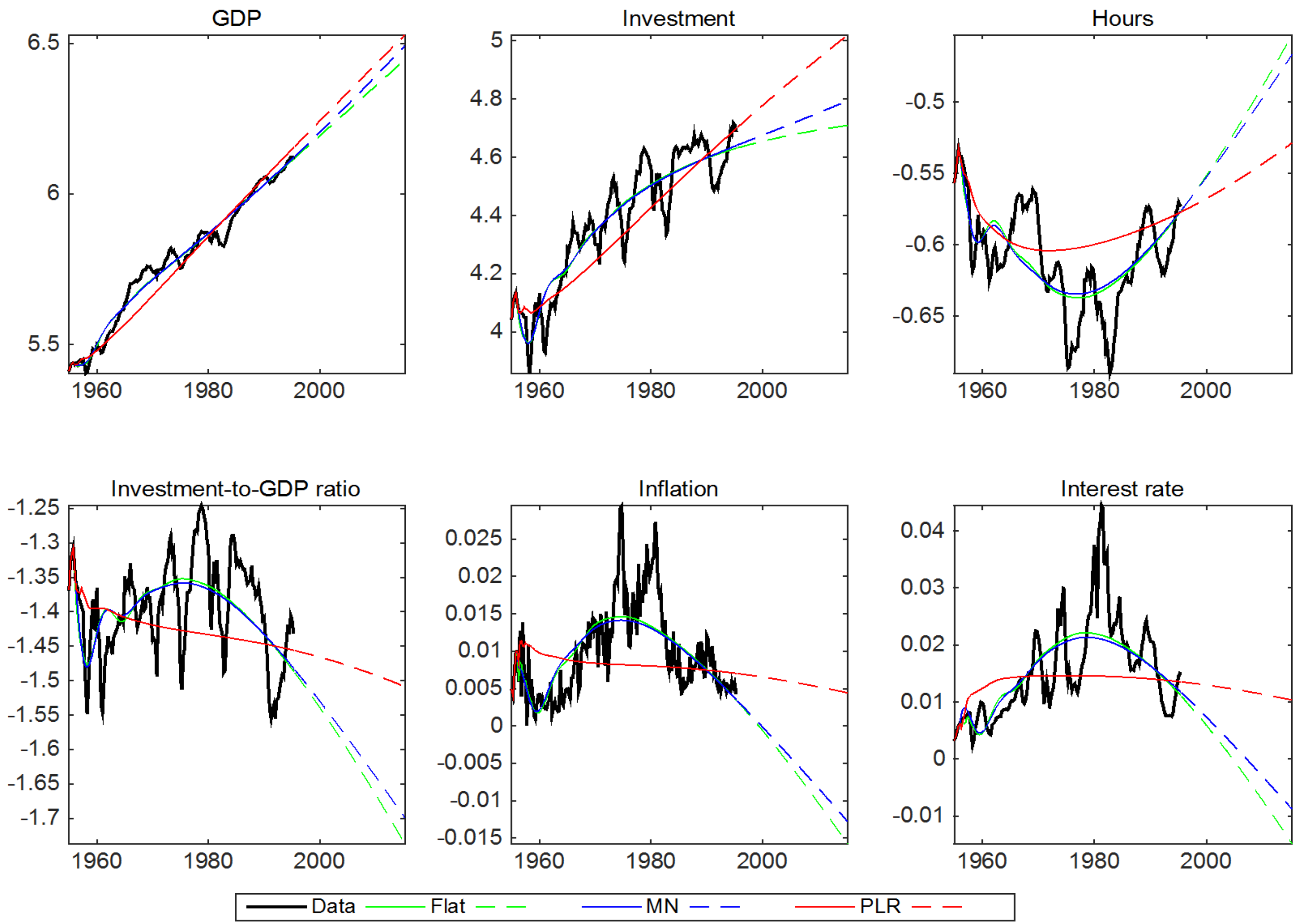
Interpretation of ***H y***

- Real trend
- Consumption-to-GDP ratio
- Investment-to-GDP ratio
- Labor share
- Hours
- Real interest rate
- Nominal trend

Deterministic components in VARs



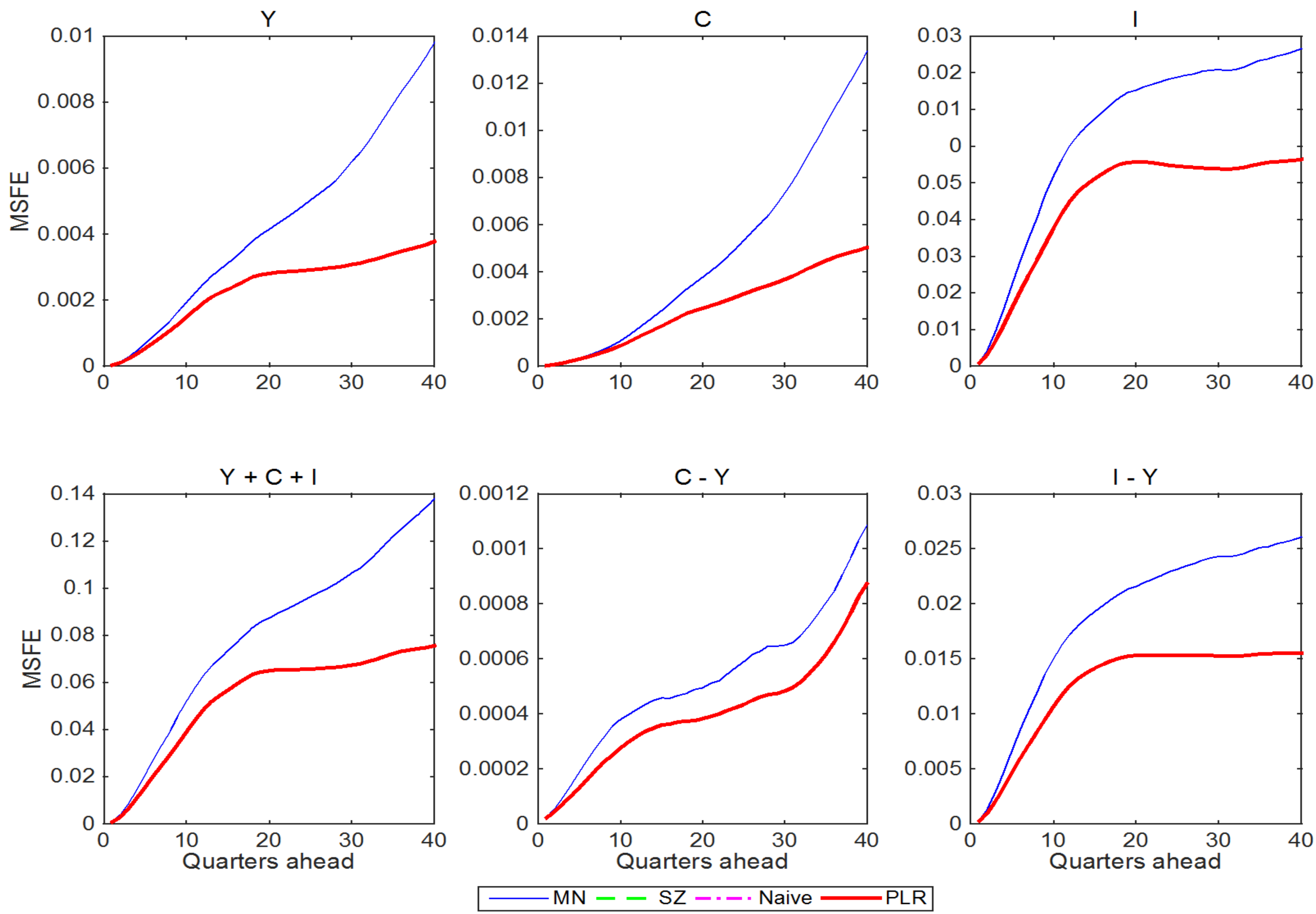
Deterministic components in VARs with Prior for the Long Run



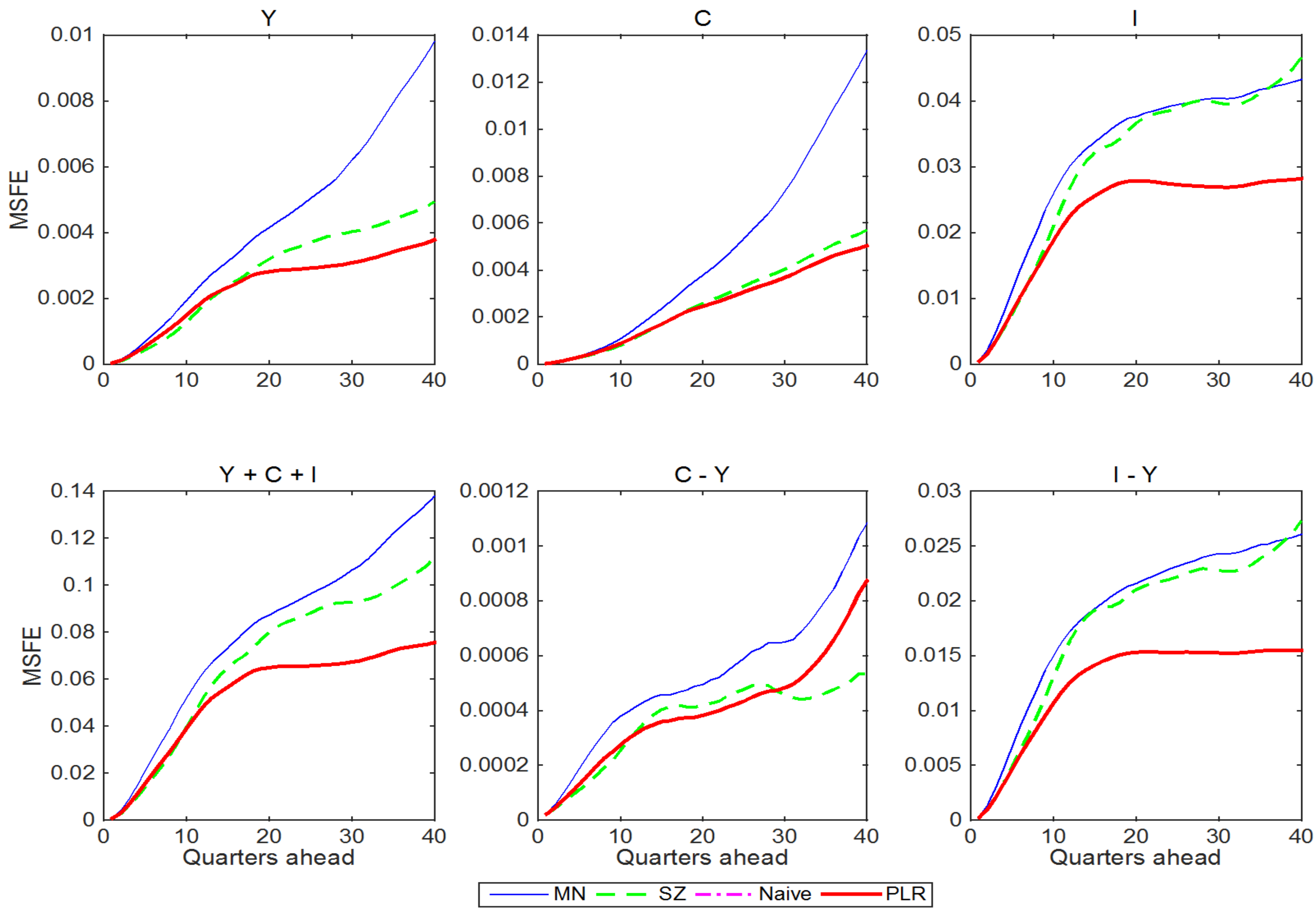
Forecasting results with 3-, 5- and 7-variable VARs

- Recursive estimation starts in 1955:I
- Forecast-evaluation sample: 1985:I – 2013:I

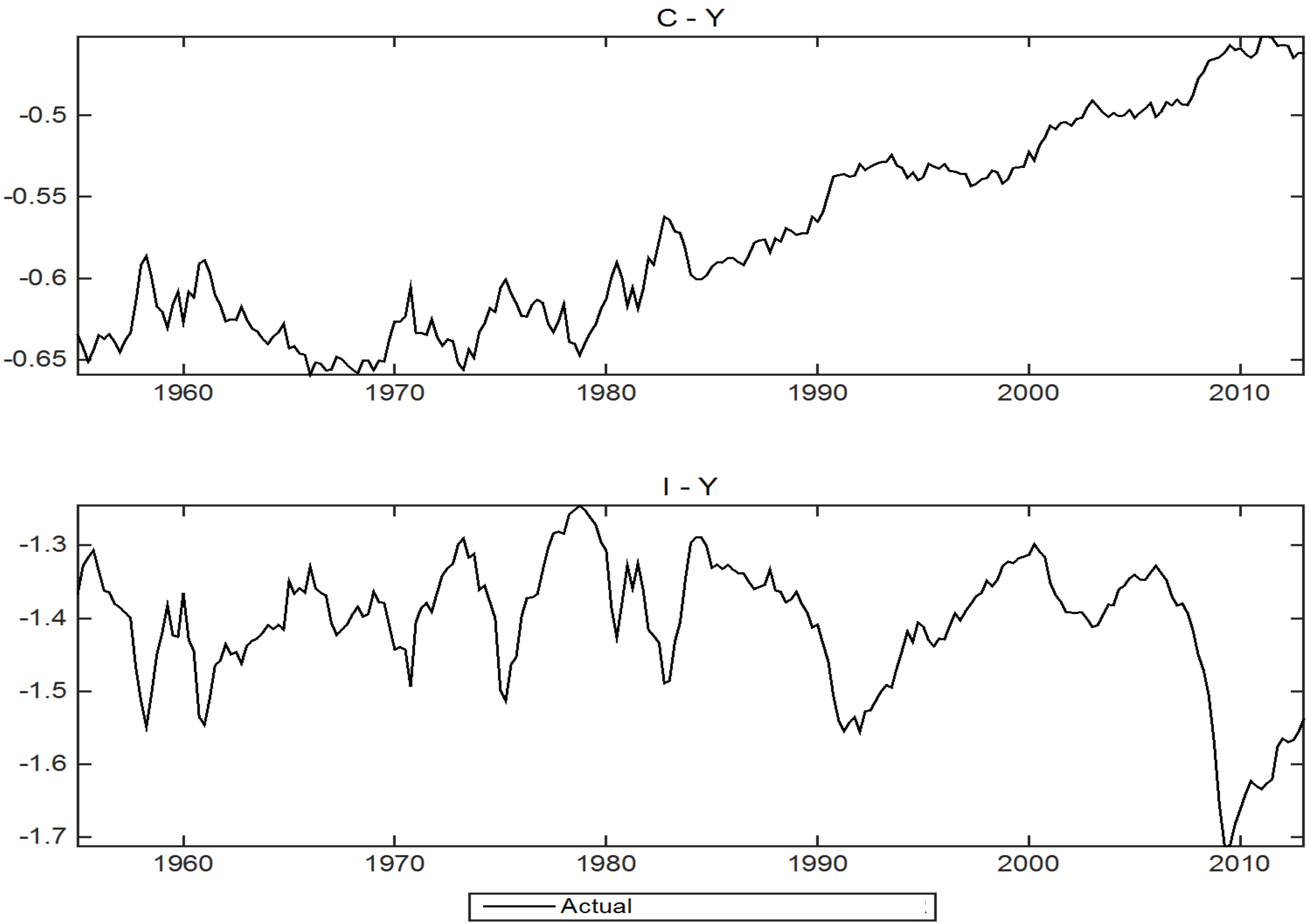
3-variable VAR: MSFE (1985-2013)



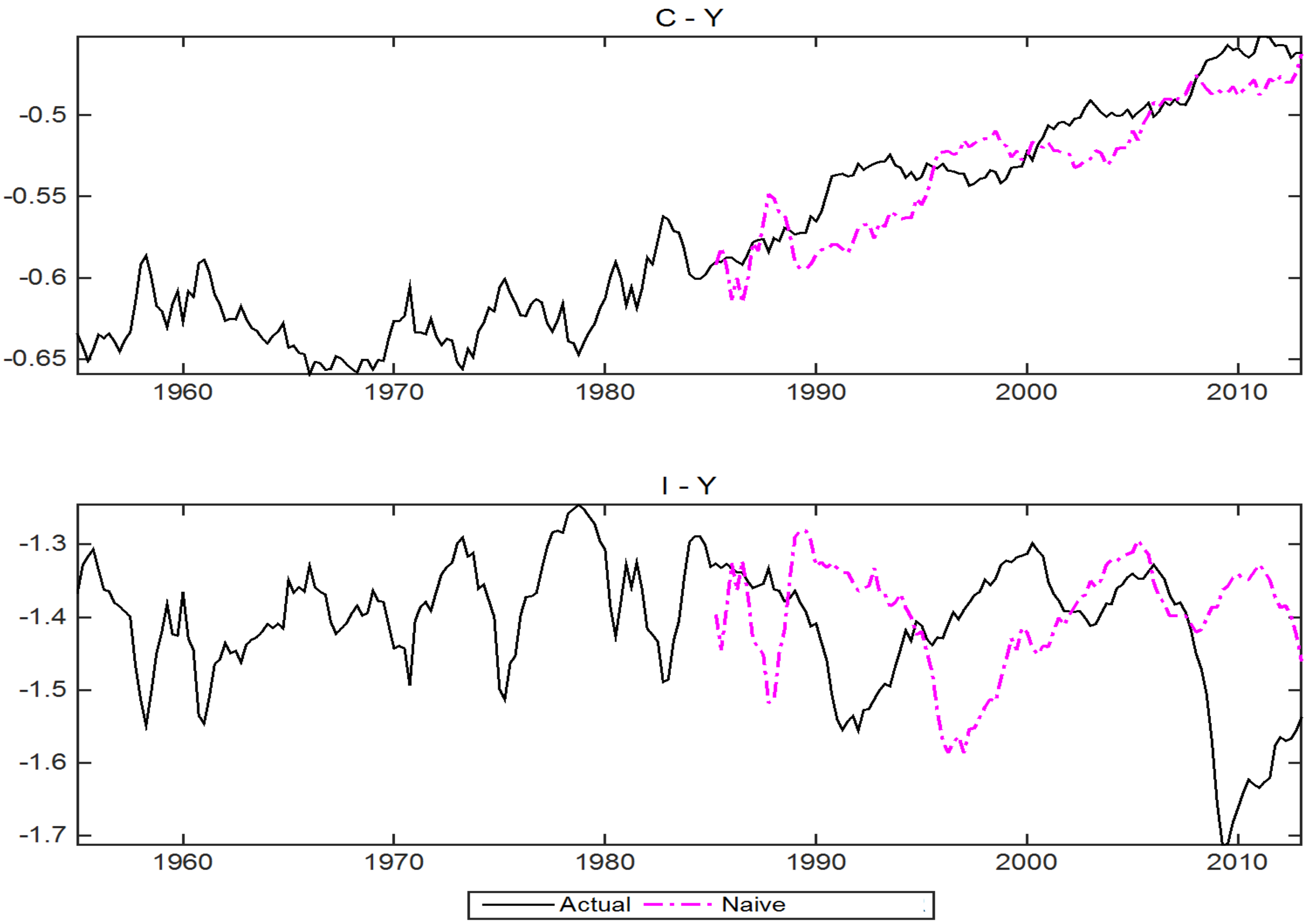
3-variable VAR: MSFE (1985-2013)



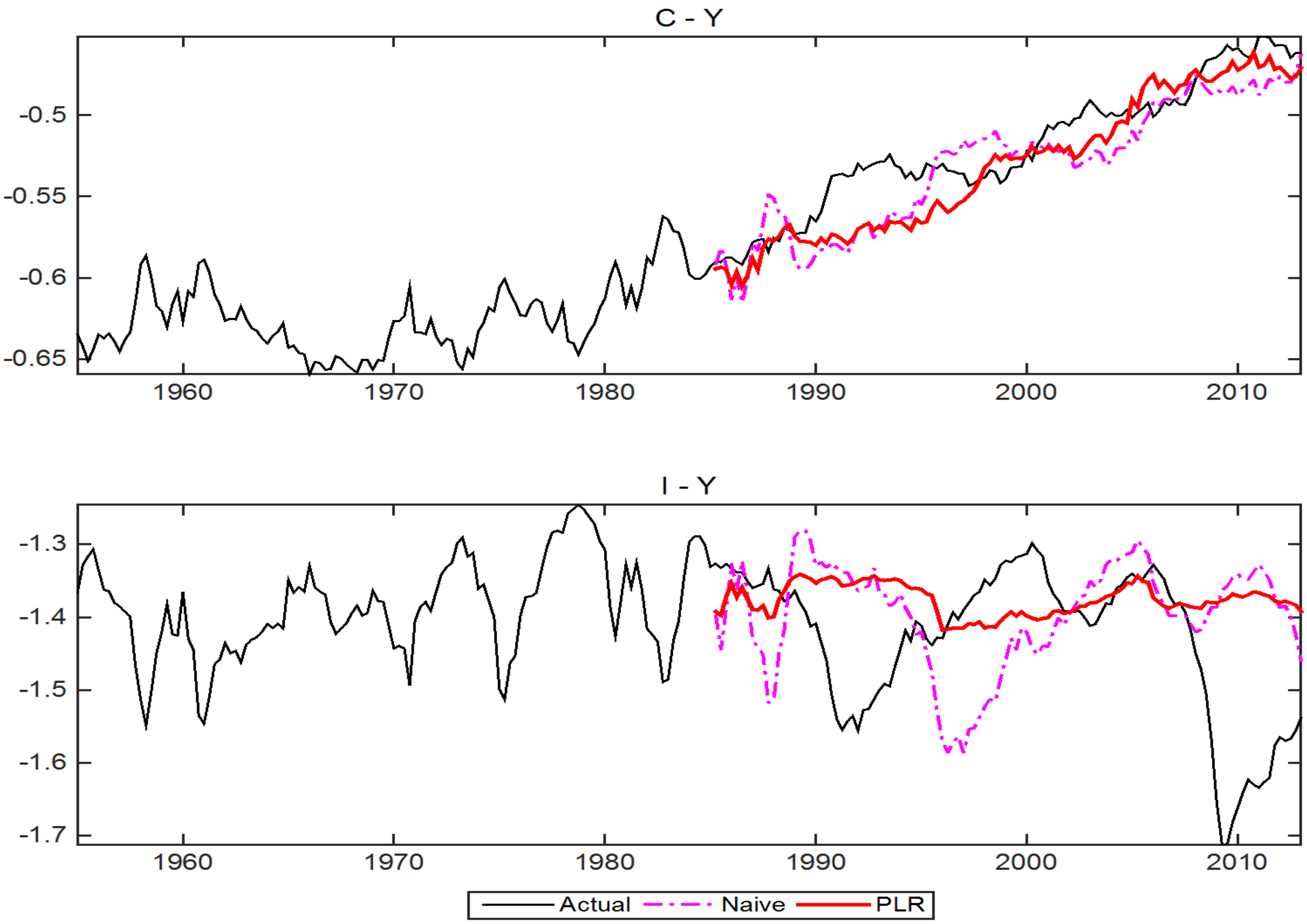
Consumption- and Investment-to-GDP ratios



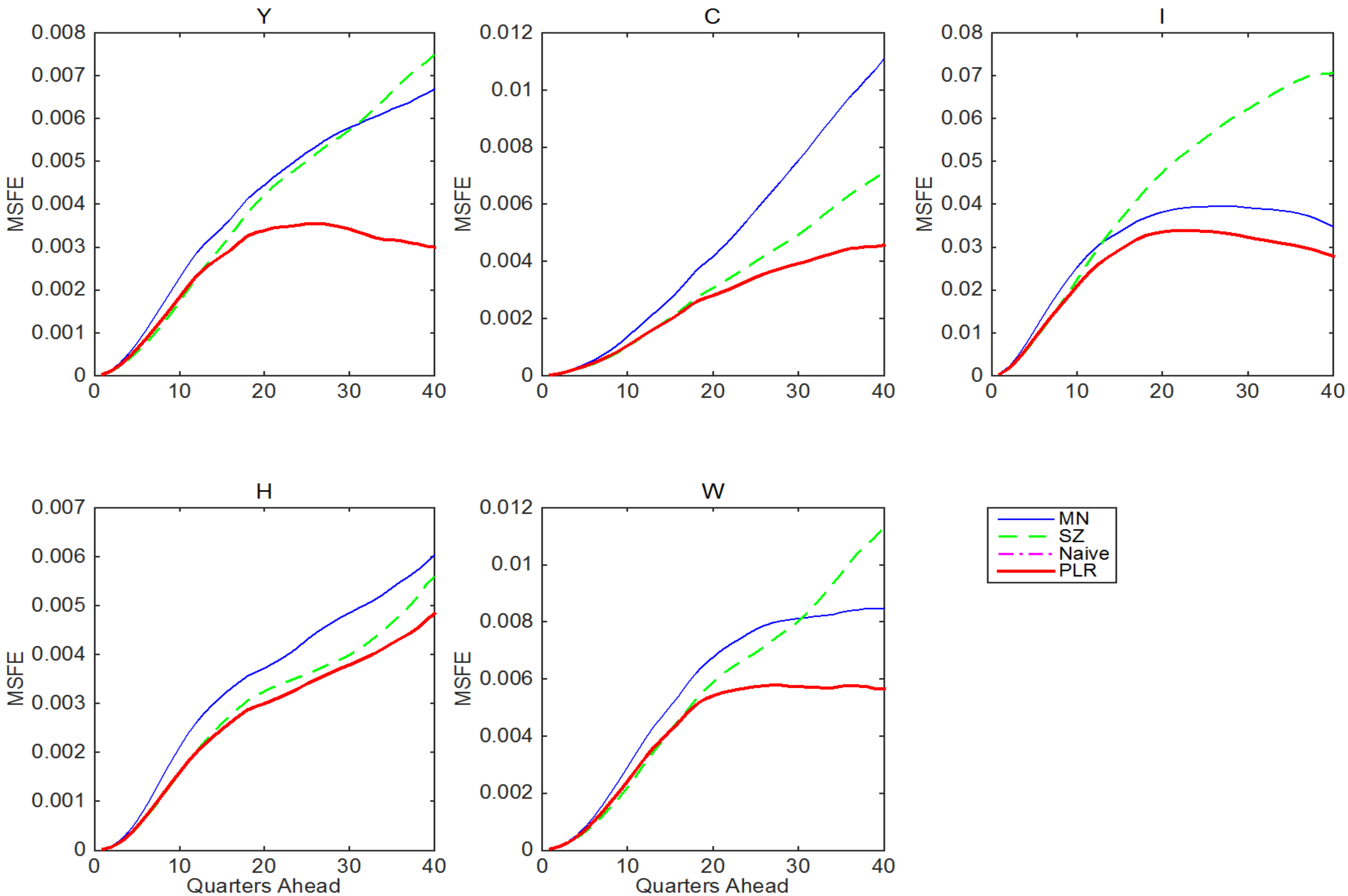
Forecasts (5 years ahead)



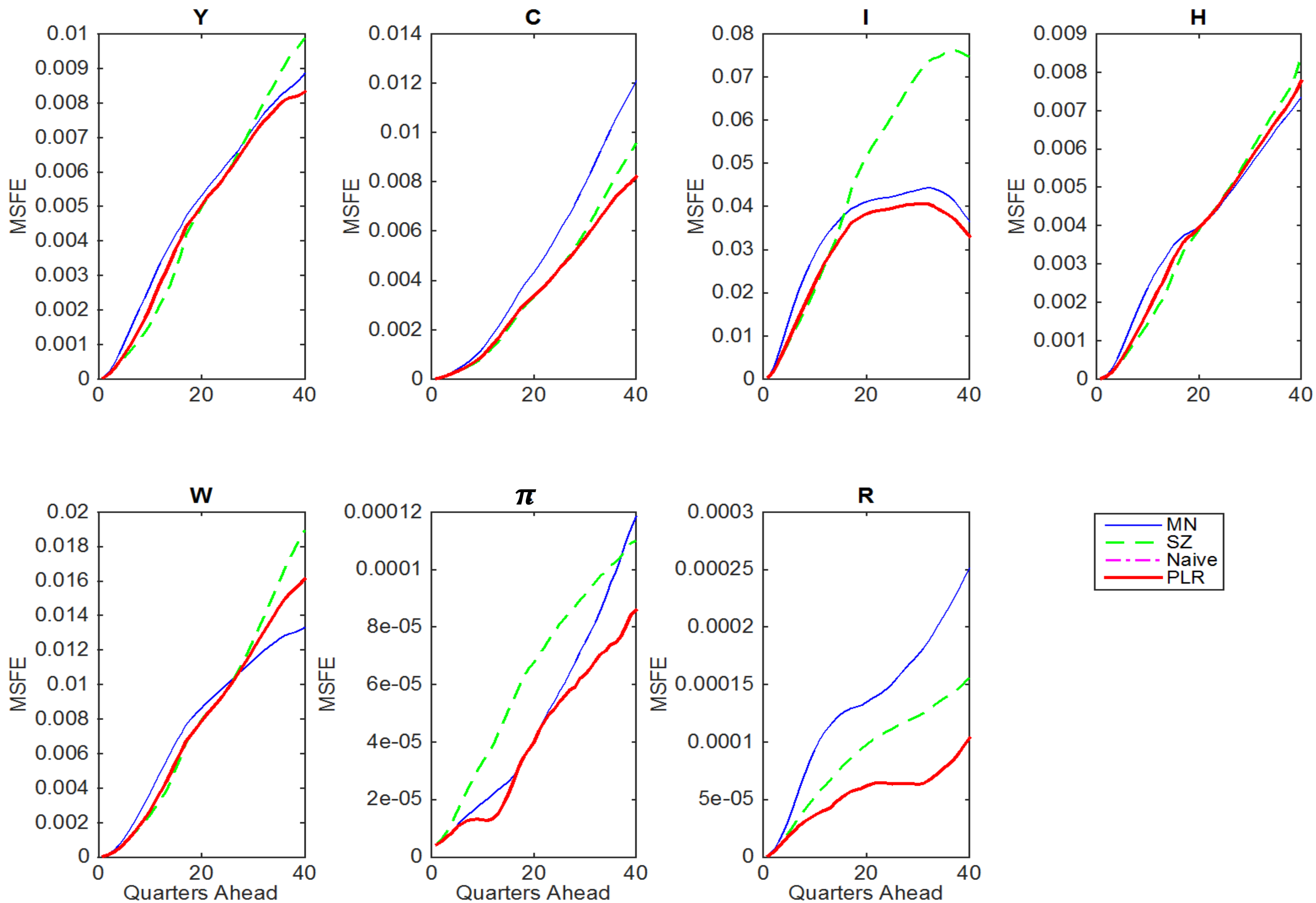
Forecasts (5 years ahead)



5-variable VAR: MSFE (1985-2013)



7-variable VAR: MSFE (1985-2013)



Invariance to rotations of the “stationary” space

- Our baseline prior depends on the choice of a specific ***H*** matrix

$$H = \begin{bmatrix} \beta_\perp' \\ \beta' \end{bmatrix}$$

Invariance to rotations of the "stationary" space

- Our baseline prior depends on the choice of a specific ***H*** matrix

$$H = \begin{bmatrix} \beta'_{\perp} \\ \beta' \end{bmatrix}$$

- Economic theory is useful, but not sufficient to uniquely pin down ***H***
  - Macro models are typically informative about $\boldsymbol{\beta}_{\perp}$ and ***sp(β)***

Invariance to rotations of the "stationary" space

- Our baseline prior depends on the choice of a specific ***H*** matrix

$$H = \begin{bmatrix} \beta_\perp' \\ \beta' \end{bmatrix}$$

- Economic theory is useful, but not sufficient to uniquely pin down ***H***
  - Macro models are typically informative about $\boldsymbol{\beta}_\perp$ and ***sp(β)***

➠ Extension of our PLR that is invariant to rotations of $\boldsymbol{\beta}$

Invariance to rotations of the "stationary" space

- Our baseline prior depends on the choice of a specific ***H*** matrix

$$H = \begin{bmatrix} \beta_\perp' \\ \beta' \end{bmatrix}$$

- Economic theory is useful, but not sufficient to uniquely pin down ***H***
  - Macro models are typically informative about $\boldsymbol{\beta}_\perp$ and ***sp(β)***

➠ Extension of our PLR that is invariant to rotations of $\boldsymbol{\beta}$

Baseline PLR: $\Lambda_{\cdot i} \cdot (H_{i\cdot}\bar{y}_0)|H,\Sigma \sim N\left(0, \phi_i^2\Sigma\right), \quad i = 1, \dots, n$

Invariance to rotations of the "stationary" space

- Our baseline prior depends on the choice of a specific ***H*** matrix

$$H = \begin{bmatrix} \beta_\perp' \\ \beta' \end{bmatrix}$$

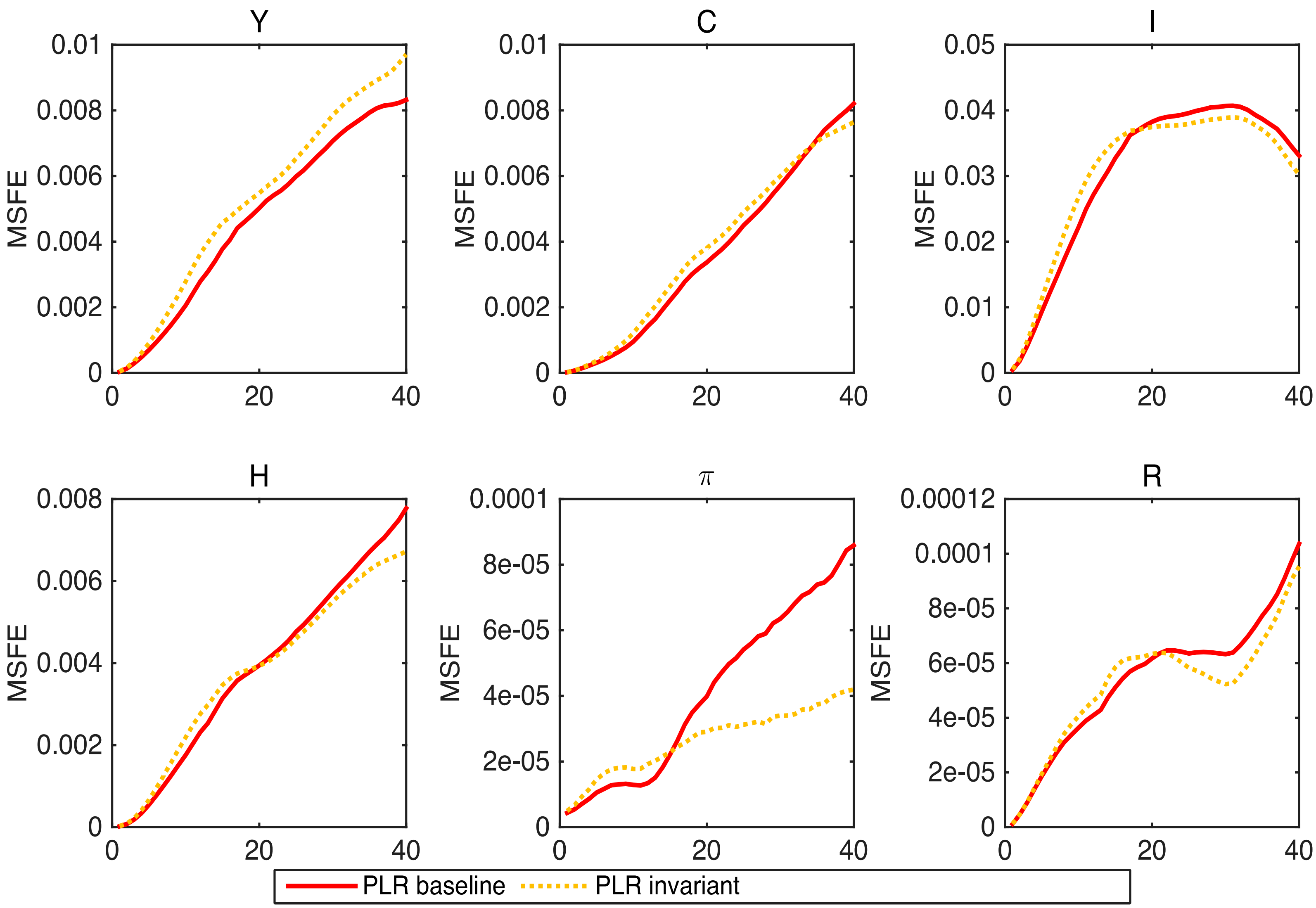
- Economic theory is useful, but not sufficient to uniquely pin down ***H***
  - Macro models are typically informative about $\boldsymbol{\beta}_\perp$ and ***sp(β)***

➠ Extension of our PLR that is invariant to rotations of ***β***

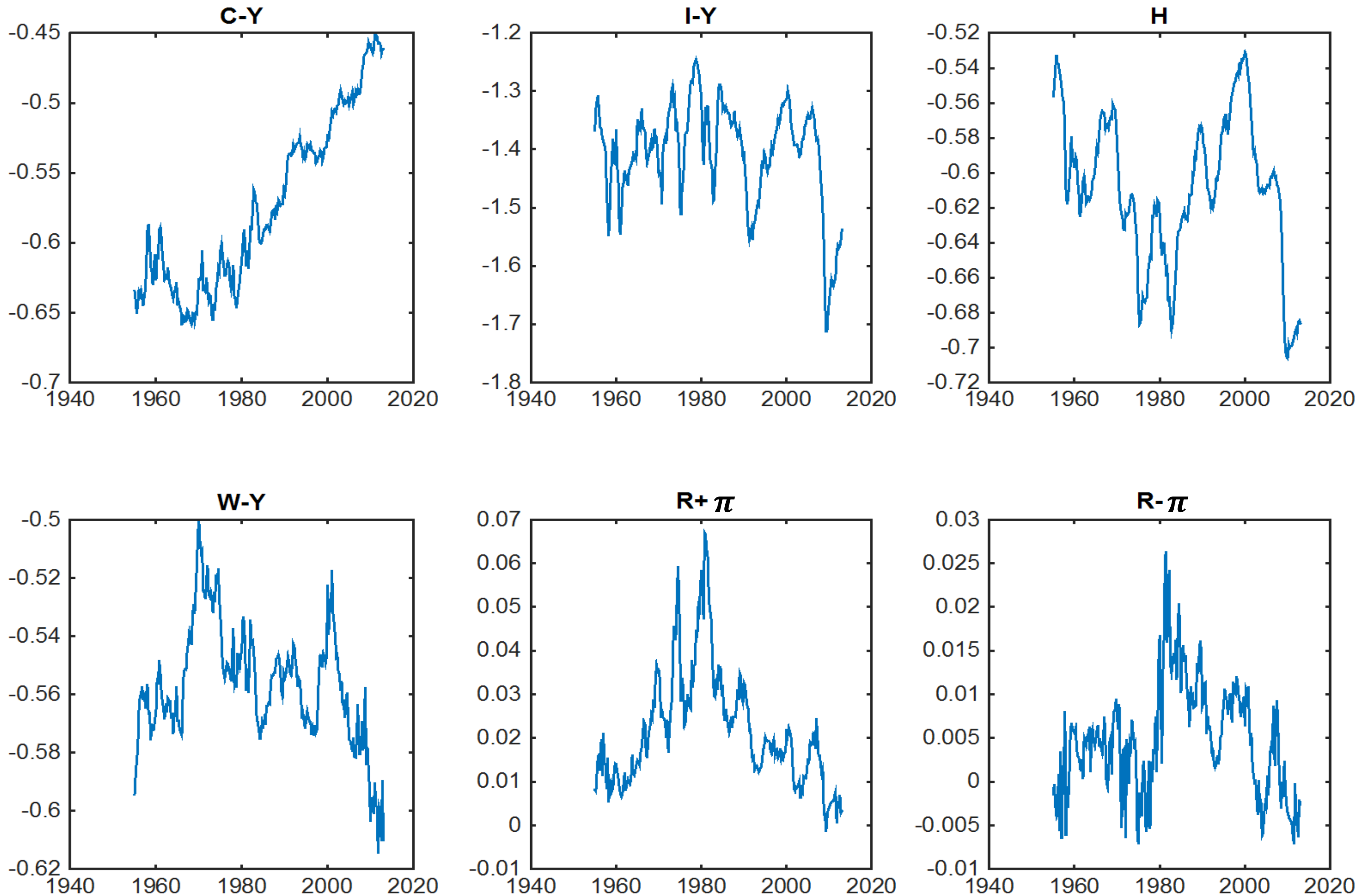
Baseline PLR: $\Lambda_{\cdot i} \cdot (H_{i\cdot}\bar{y}_0) | H, \Sigma \sim N\left(0, \phi_i^2 \Sigma\right), \quad i = 1, \dots, n$

Invariant PLR:
$$\begin{cases} \Lambda_{\cdot i} \cdot (H_{i\cdot}\bar{y}_0) | H, \Sigma \sim N\left(0, \phi_i^2 \Sigma\right), \quad i = 1, \dots, n-r \\ \sum_{i=n-r+1}^{n} \Lambda_{\cdot i} \cdot (H_{i\cdot}\bar{y}_0) | H, \Sigma \sim N(0, \phi_{n-r+1}^2 \Sigma) \end{cases}$$

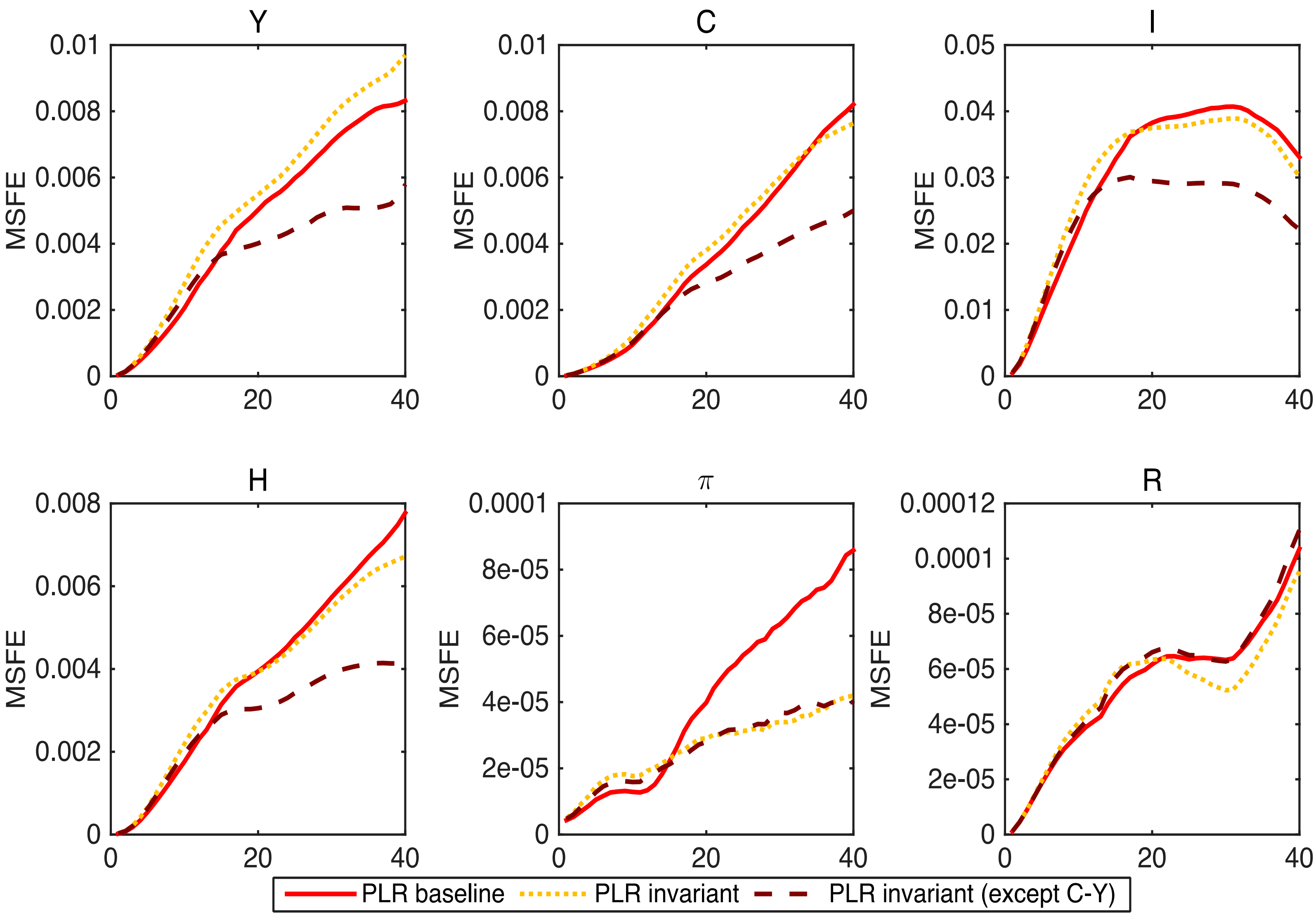
7-variable VAR: Forecasting results with "invariant" PLR



$H\,y$ in the data



7-variable VAR: Forecasting results with “invariant” PLR



Strengths and weaknesses

- Strengths
  - Imposes discipline on long-run behavior of the model
  - Based on robust lessons of theoretical macro models
  - Performs well in forecasting (especially at longer horizons)
  - Very easy to implement

Strengths and weaknesses

- Strengths
  - Imposes discipline on long-run behavior of the model
  - Based on robust lessons of theoretical macro models
  - Performs well in forecasting (especially at longer horizons)
  - Very easy to implement

- “Weak” points
  - Non-automatic procedure → need to think about it
  - Might prove difficult to set up in large-scale models → might require too much thinking

Connections and extreme cases

$$\Delta y_t = c + \underbrace{\Pi \; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- Rewrite as

$$\Delta y_t = c + \begin{bmatrix} \Lambda_1 & \Lambda_2 \end{bmatrix} \begin{bmatrix} \beta_\perp ' \\ \beta ' \end{bmatrix} y_{t-1} + \varepsilon_t$$

Connections and extreme cases

$$\Delta y_t = c + \underbrace{\Pi \;\; H^{-1}}_{\Lambda} \; \underbrace{H y_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$

- Rewrite as

$$\Delta y_t = c + \begin{bmatrix} \Lambda_1 & \Lambda_2 \end{bmatrix} \begin{bmatrix} \beta_\perp ' \\ \beta' \end{bmatrix} y_{t-1} + \varepsilon_t$$

$$\Delta y_t = c + \Lambda_1 \beta_\perp ' y_{t-1} + \Lambda_2 \beta' y_{t-1} + \varepsilon_t$$

Connections and extreme cases

$$\Delta y_t = c + \Lambda_1 \beta_\perp' y_{t-1} + \Lambda_2 \beta' y_{t-1} + \varepsilon_t$$

Connections and extreme cases

$$\Delta y_t = c + \Lambda_1 \beta_\perp' y_{t-1} + \Lambda_2 \beta' y_{t-1} + \varepsilon_t$$

- Error Correction Model: dogmatic prior on $\boldsymbol{\Lambda_1} = 0$

Connections and extreme cases

$$\Delta y_t = c + \Lambda_1 \beta_\perp' y_{t-1} + \Lambda_2 \beta' y_{t-1} + \varepsilon_t$$

- Error Correction Model: dogmatic prior on $\boldsymbol{\Lambda_1} = 0$

- KPSW, CEE
  - fix **β** based on theory
  - flat prior on $\boldsymbol{\Lambda_2}$

Connections and extreme cases

$$\Delta y_t = c + \Lambda_1 \beta_\perp' y_{t-1} + \Lambda_2 \beta' y_{t-1} + \varepsilon_t$$

- Error Correction Model: dogmatic prior on $\boldsymbol{\Lambda_1} = 0$

- KPSW, CEE
  - fix **β** based on theory
  - flat prior on $\boldsymbol{\Lambda_2}$
- Cointegration
  - estimate **β**
  - flat prior on $\boldsymbol{\Lambda_2}$
  - EG (1987)

Connections and extreme cases

$$\Delta y_t = c + \Lambda_1 \beta_\perp' y_{t-1} + \Lambda_2 \beta' y_{t-1} + \varepsilon_t$$

- Error Correction Model: dogmatic prior on $\boldsymbol{\Lambda_1}$ = 0

- KPSW, CEE
  - fix **β** based on theory
  - flat prior on $\boldsymbol{\Lambda_2}$
- Cointegration
  - estimate **β**
  - flat prior on $\boldsymbol{\Lambda_2}$
  - EG (1987)
- Bayesian cointegration
  - uniform prior on sp(**β**)
  - KSvDV (2006)

Connections and extreme cases

$$\Delta y_t = c + \Lambda_1 \beta_\perp' y_{t-1} + \Lambda_2 \beta' y_{t-1} + \varepsilon_t$$

- Error Correction Model: dogmatic prior on $\boldsymbol{\Lambda_1} = 0$

- KPSW, CEE
  - fix **β** based on theory
  - flat prior on $\boldsymbol{\Lambda_2}$
- Cointegration
  - estimate **β**
  - flat prior on $\boldsymbol{\Lambda_2}$
  - EG (1987)
- Bayesian cointegration
  - uniform prior on sp(**β**)
  - KSvDV (2006)

- VAR in first differences: dogmatic prior on $\boldsymbol{\Lambda_1} = \boldsymbol{\Lambda_2} = 0$

Connections and extreme cases

$$\Delta y_t = c + \Lambda_1 \beta_\perp' y_{t-1} + \Lambda_2 \beta' y_{t-1} + \varepsilon_t$$

- Error Correction Model: dogmatic prior on $\Lambda_1 = 0$

- KPSW, CEE
  - fix **β** based on theory
  - flat prior on Λ_2
- Cointegration
  - estimate **β**
  - flat prior on Λ_2
  - EG (1987)
- Bayesian cointegration
  - uniform prior on sp(**β**)
  - KSvDV (2006)

- VAR in first differences: dogmatic prior on $\Lambda_1 = \Lambda_2 = 0$

- Sum-of-coefficients prior (DLS, SZ)
  - $[\beta' \;\; \beta']' = H = I$
  - shrink Λ_1 and Λ_2 to 0

3-var VAR: Mean Squared Forecast Errors (1985-2013)