

Priors for the long run

Domenico Giannone

New York Fed

Michele Lenza

European Central Bank

Giorgio Primiceri

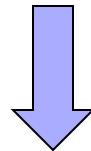
Northwestern University

2015 NBER-NSF Time Series Conference

Vienna, September 25-26

What we do

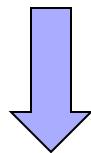
- Propose a class of prior distributions for VARs that discipline the long-run implications of the model



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 → Trivial to implement
 - Can be easily combined 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 $\rho = 1$, 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 $\rho = 1$, 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 y_0
 - No penalization for parameter estimates of implying steady states or trends far away from initial conditions

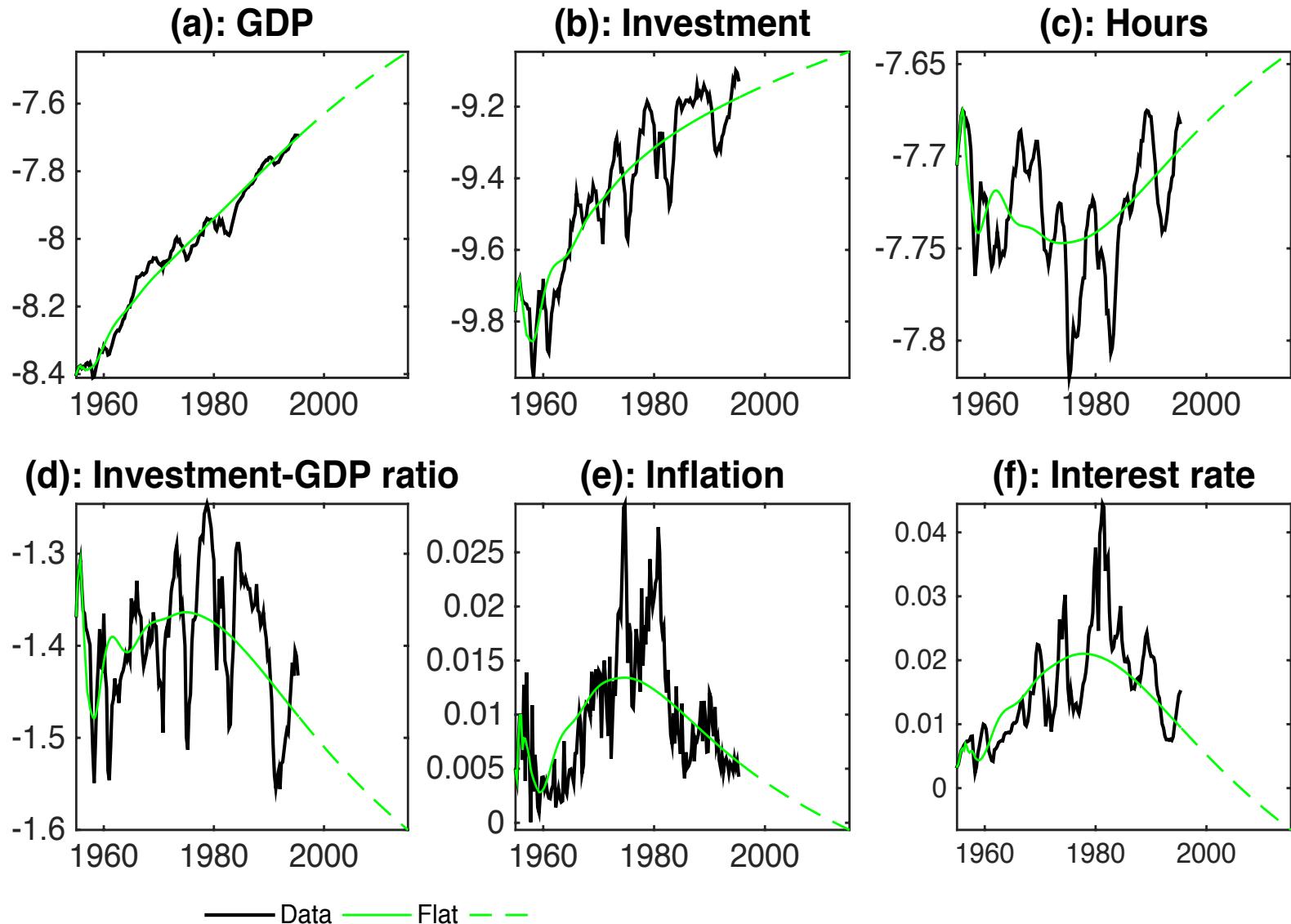
Deterministic components in VARs

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

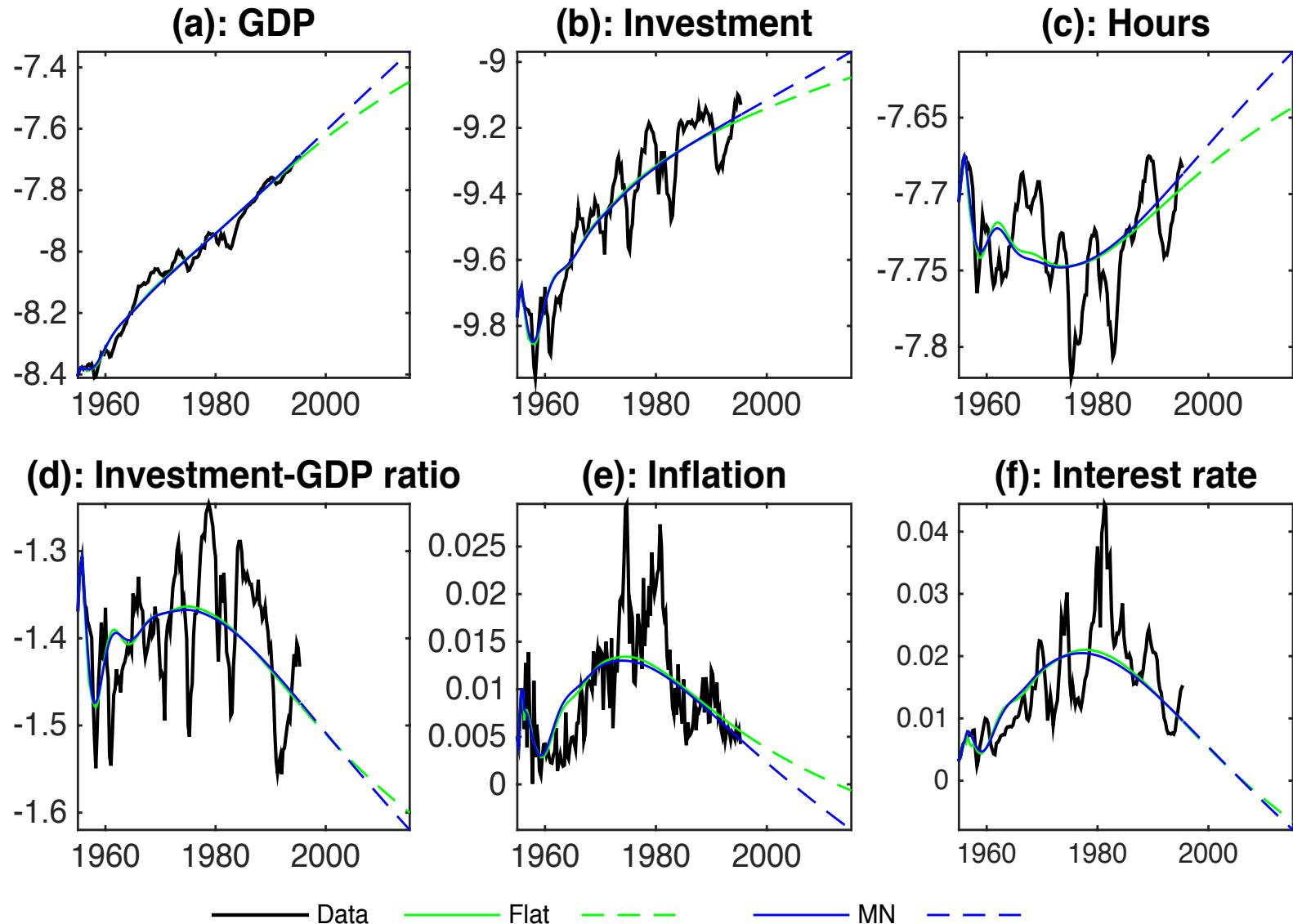
Deterministic components in VARs

- 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

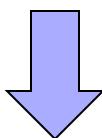


Pathology of (flat-prior) VARs (Sims, 1996 and 2000)

- 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 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 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

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

Prior for the long run

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

- 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

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

- 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

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

- 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

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

Prior for the long run

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

- Rewrite as

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

- 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

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

- 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

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

- 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

$$\Delta y_t = c + \underbrace{\prod_{\Lambda} H^{-1}}_{\tilde{y}_{t-1}} \underbrace{Hy_{t-1}}_{\tilde{y}_{t-1}} + \varepsilon_t$$
$$\left[\begin{array}{ccc} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{array} \right] \left[\begin{array}{c} \text{Output} \\ \text{Consumption} \\ \text{Investment} \end{array} \right]$$

Example: 3-variable VAR of KPSW

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

$\left[\begin{array}{ccc} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{array} \right]$

 Output
 Consumption
 Investment

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

Example: 3-variable VAR of KPSW

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

$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$
Output
Consumption
Investment

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

c_{t-1} - x_{t-1}
i_{t-1} - x_{t-1}

←
Possibly stationary linear combinations

Example: 3-variable VAR of KPSW

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

$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$
Output
Consumption
Investment

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

x_{t-1} + c_{t-1} + i_{t-1}
c_{t-1} - x_{t-1}
i_{t-1} - x_{t-1}

Common trend
←

Possibly stationary linear combinations
←

Connections and extreme cases

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

- Rewrite as

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

Connections and extreme cases

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

- Rewrite as

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

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

Connections and extreme cases

$$\Delta y_t = c + \Lambda_1 \beta_1' y_{t-1} + \Lambda_2 \beta_2' y_{t-1} + \varepsilon_t$$

Connections and extreme cases

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

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

Connections and extreme cases

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

- Error Correction Model: dogmatic prior on $\Lambda_1 = 0$
- KPSW, CEE
 - fix β based on theory
 - flat prior on Λ_2

Connections and extreme cases

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

- 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)

Connections and extreme cases

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

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

- | | | |
|---|--|--|
| ➤ KPSW, CEE <ul style="list-style-type: none">■ fix β based on theory■ flat prior on Λ_2 | ➤ Cointegration <ul style="list-style-type: none">■ estimate β■ flat prior on Λ_2■ EG (1987) | ➤ Bayesian cointegration <ul style="list-style-type: none">■ uniform prior on $\text{sp}(\beta)$■ KSvDV (2006) |
|---|--|--|

Connections and extreme cases

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

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

- | | | |
|---|--|---|
| ➤ KPSW, CEE | ➤ Cointegration | ➤ Bayesian cointegration |
| <ul style="list-style-type: none">■ fix β based on theory■ flat prior on Λ_2 | <ul style="list-style-type: none">■ estimate β■ flat prior on Λ_2■ EG (1987) | <ul style="list-style-type: none">■ uniform prior on $\text{sp}(\beta)$■ KSvDV (2006) |

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

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 | ➤ Cointegration | ➤ Bayesian cointegration |
| <ul style="list-style-type: none">■ fix β based on theory■ flat prior on Λ_2 | <ul style="list-style-type: none">■ estimate β■ flat prior on Λ_2■ EG (1987) | <ul style="list-style-type: none">■ uniform prior on $\text{sp}(\beta)$■ KSvDV (2006) |

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

■ Sum-of-coefficients prior (DLS, SZ)

- $[\beta_{\perp}' \beta']' = H = I$
- shrink Λ_1 and Λ_2 to 0

Prior for the long run: specification and implementation

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

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

Prior for the long run: specification and implementation

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

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

Prior for the long run: specification and implementation

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

- $\Lambda_{\cdot i} | H, \Sigma \sim N \left(0, \phi_i^2 \frac{\Sigma}{(H_{\cdot i} y_0)^2} \right), \quad i = 1, \dots, n$
- Conjugate
 - Trivial to implement with Theil mixed estimation in the VAR in levels

Prior for the long run: specification and implementation

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

- $\Lambda_{\cdot i} | H, \Sigma \sim N \left(0, \phi_i^2 \frac{\Sigma}{(H_{\cdot i} y_0)^2} \right), \quad i = 1, \dots, n$
- Conjugate
 - Trivial to implement with Theil mixed estimation in the VAR in levels
 - Can be easily combined with existing priors

Prior for the long run: specification and implementation

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

$$\blacksquare \quad \Lambda_{\cdot i} | H, \Sigma \sim N \left(0, \phi_i^2 \frac{\Sigma}{(H_{\cdot i} y_0)^2} \right), \quad i = 1, \dots, n$$

■ Conjugate

- Trivial to implement 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 the setting of hyperparameters ϕ (GLP, 2013)

Empirical results

- Deterministic component in 7-variable VAR
- Forecasting
 - 3-variable VAR
 - 7-variable VAR

7-variable VAR

- VAR(5) with quarterly data on

- GDP
- Consumption
- Investment
- Real Wages
- Hours
- Inflation
- Federal funds rate

7-variable VAR

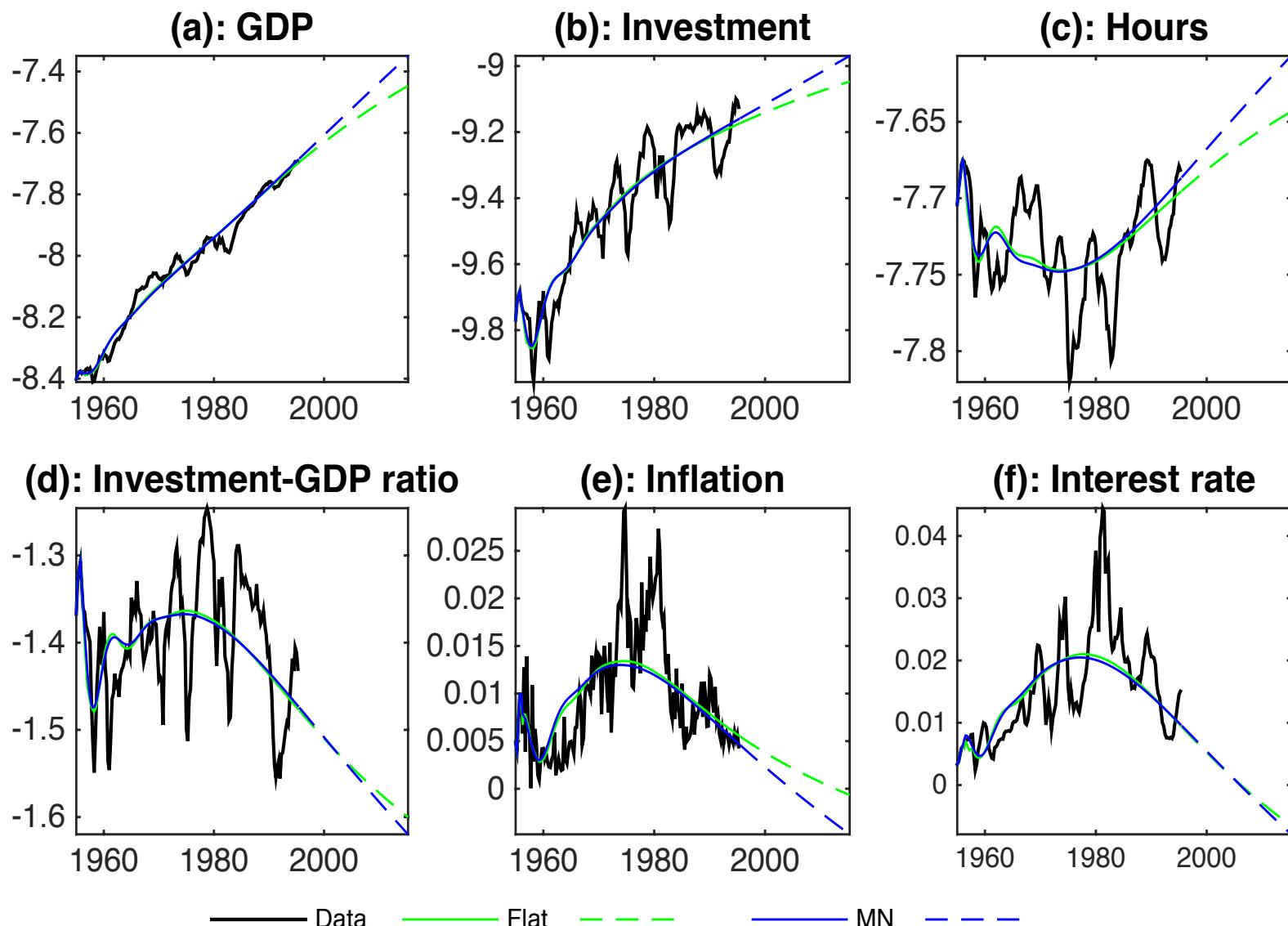
- VAR(5) with quarterly data on

- GDP
- Consumption
- Investment
- Real Wages
- Hours
- Inflation
- Federal funds rate

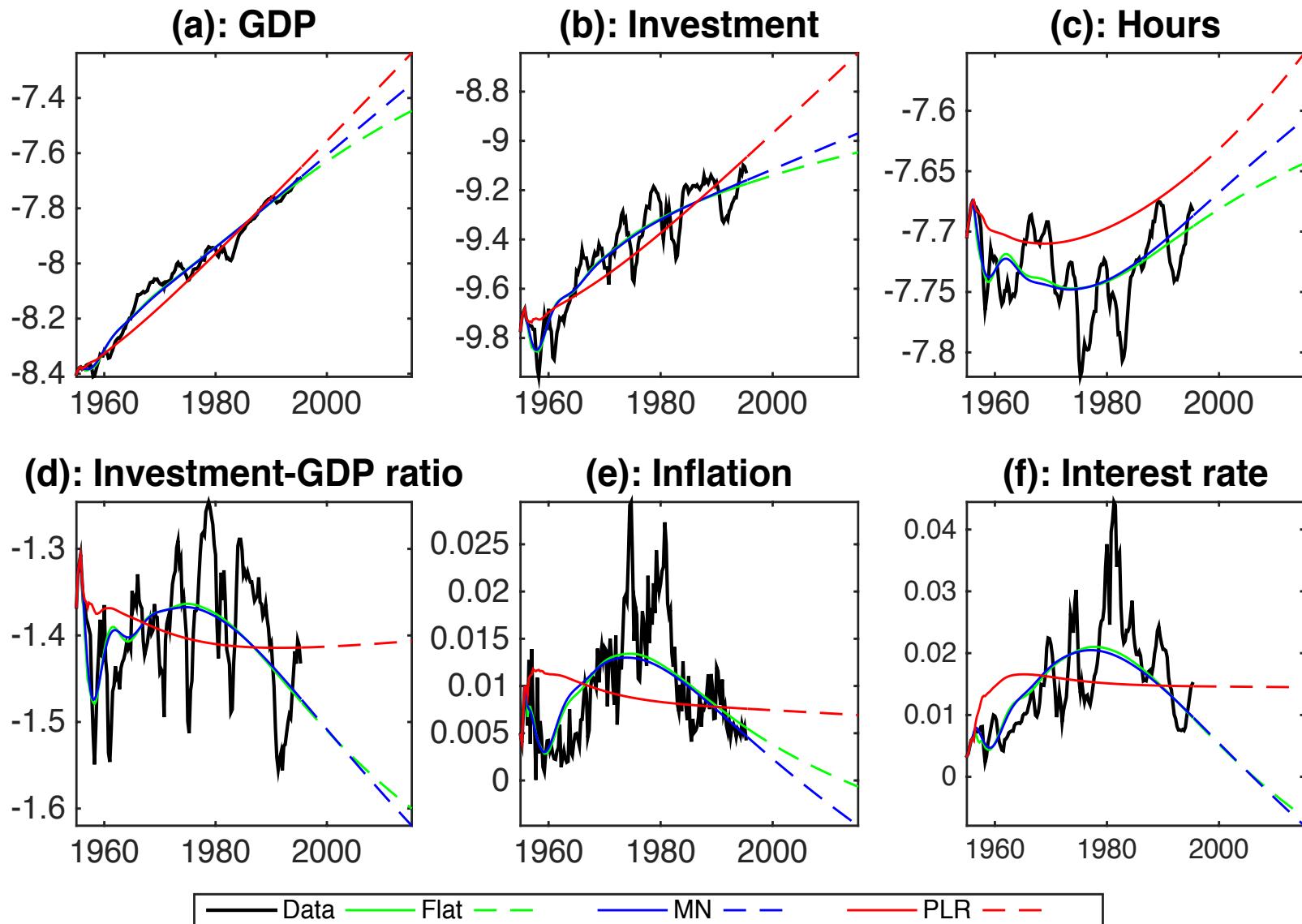
- $H = \begin{bmatrix} x & c & i & w & h & \pi & R \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$ Interpretation of $H y$

→	Real trend
→	Nominal trend
→	Cons-GDP ratio
→	Inv-GDP ratio
→	Wage-GDP ratio
→	Hours
→	Real interest rate

Deterministic components in VARs



Deterministic components in VARs with Prior for the Long Run



3-variable VAR

- VAR(5) with quarterly data on
 - GDP
 - Consumption
 - Investment
- Recursive estimation starts in 1955:I
- Forecast-evaluation sample: 1985:I – 2013:I

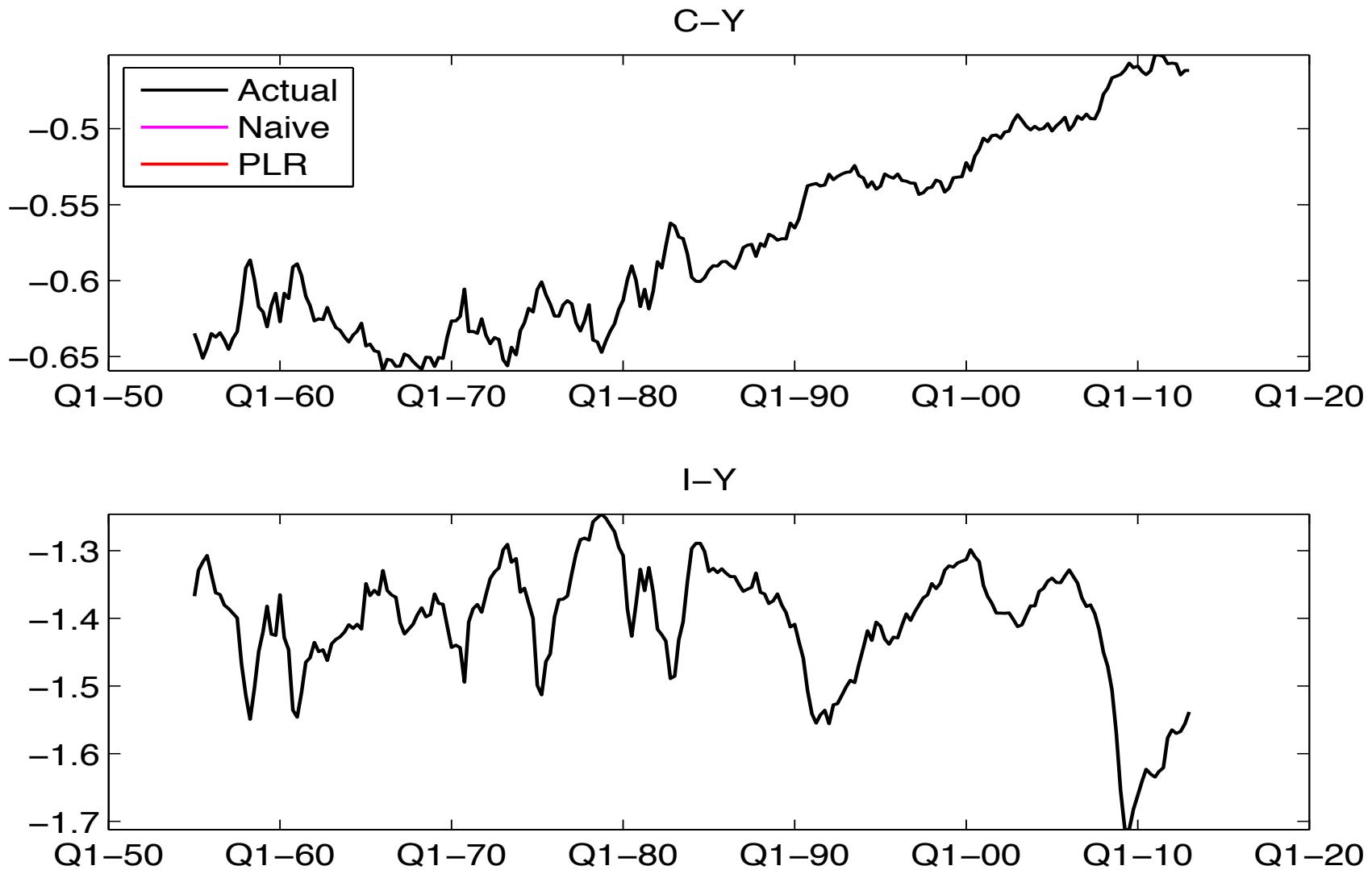
■ $H = \begin{bmatrix} x & c & i \\ 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad \xrightarrow{\hspace{1cm}} \quad \text{Interpretation of } H y$

Real trend

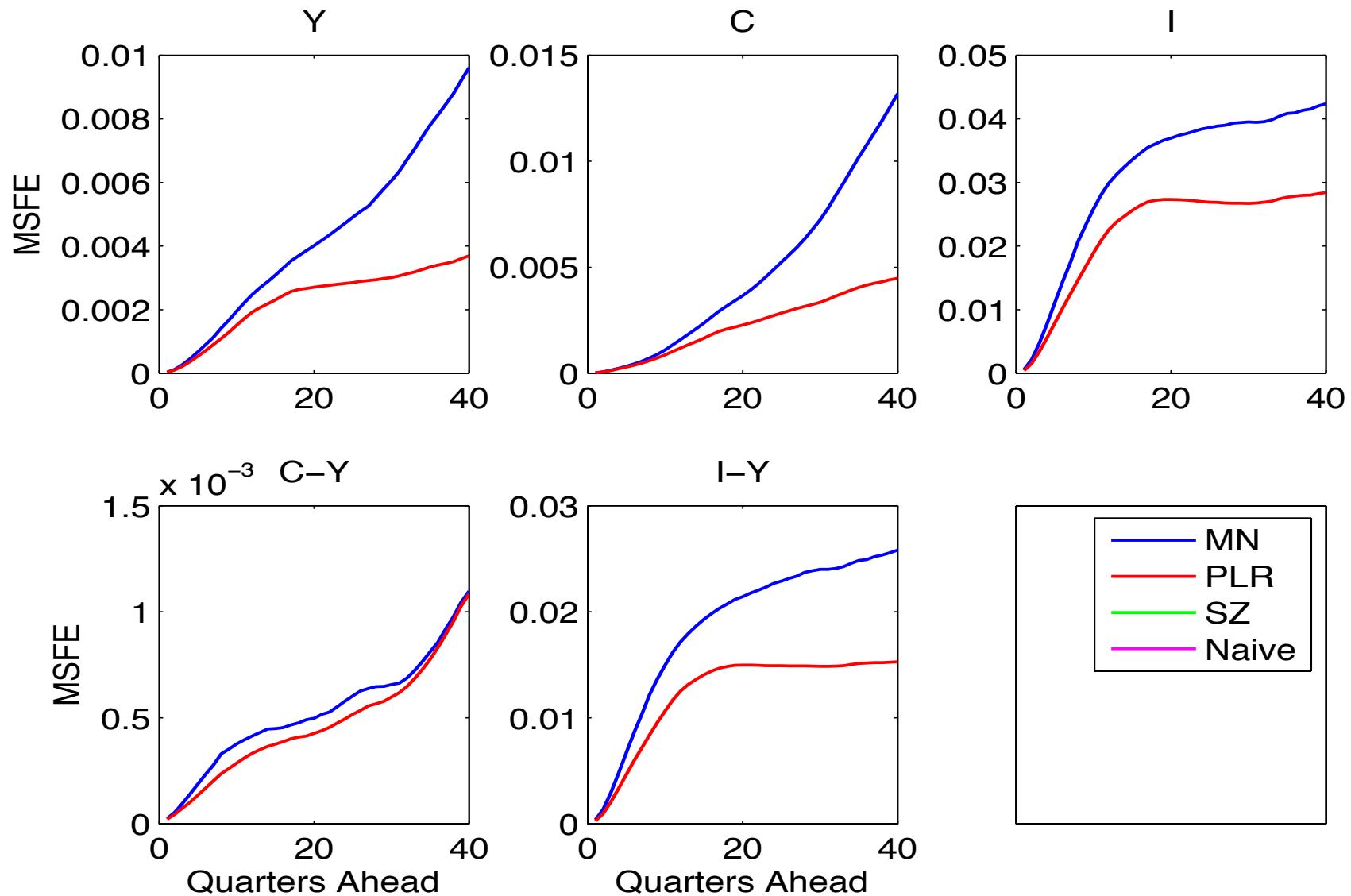
Cons-GDP ratio

Inv-GDP ratio

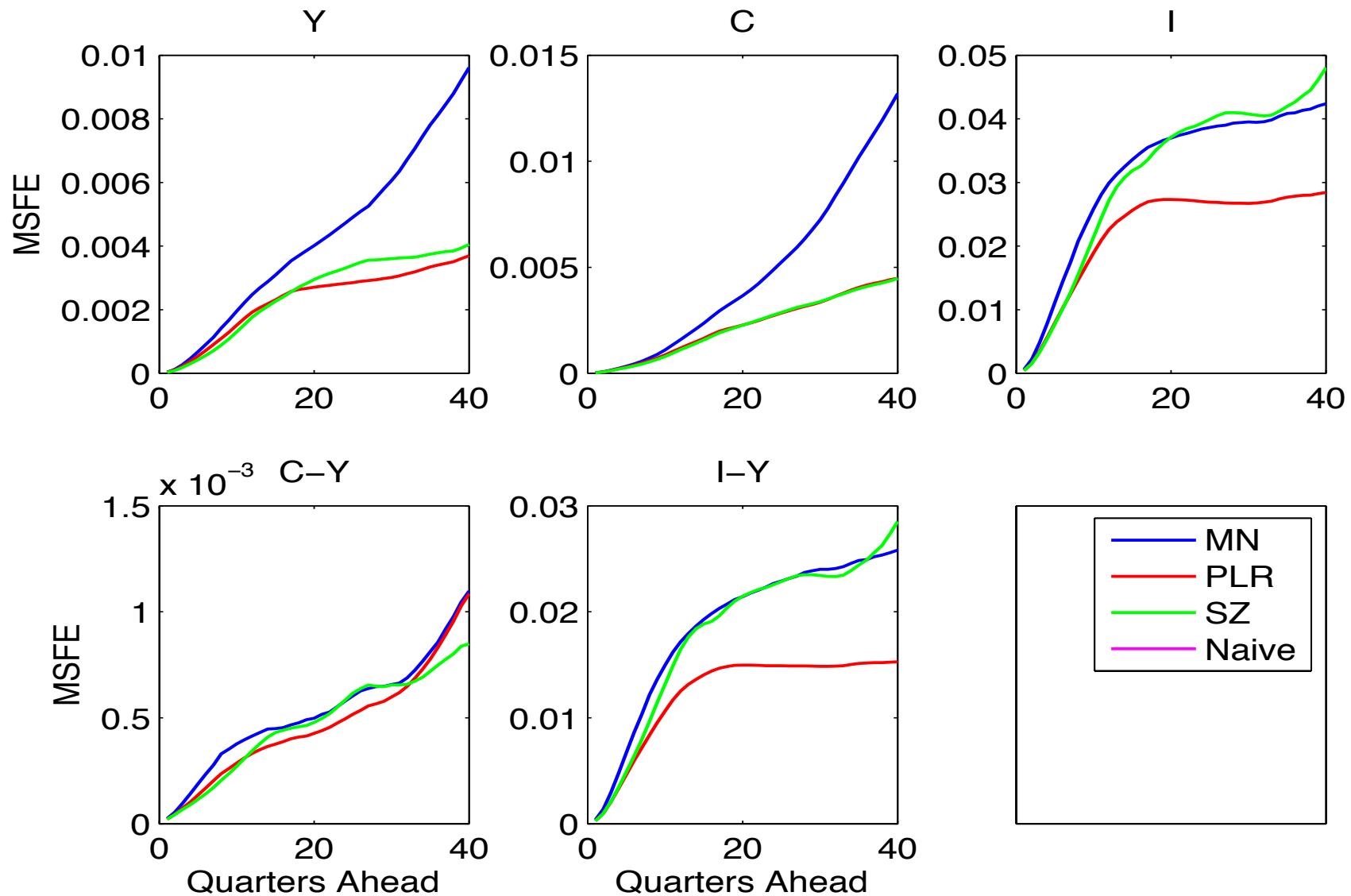
Consumption and investment-output ratios



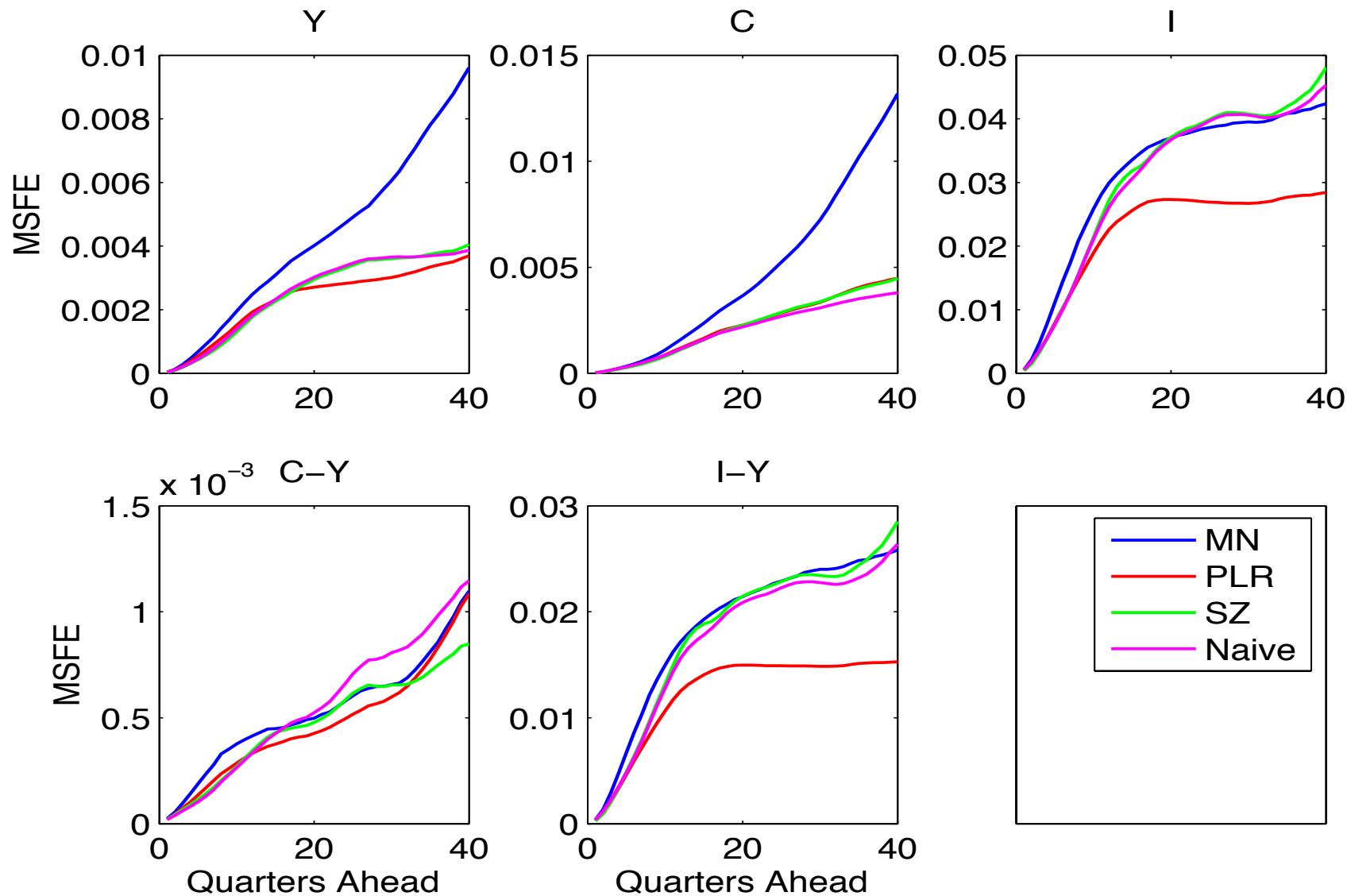
Mean Squared Forecast Errors (1985-2013)



Mean Squared Forecast Errors (1985-2013)

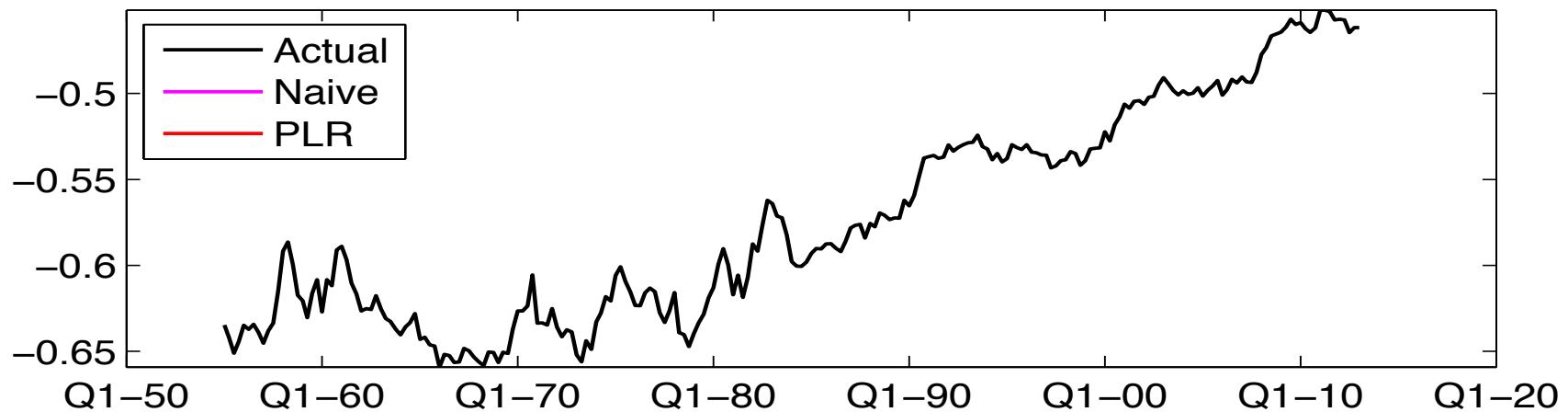


Mean Squared Forecast Errors (1985-2013)

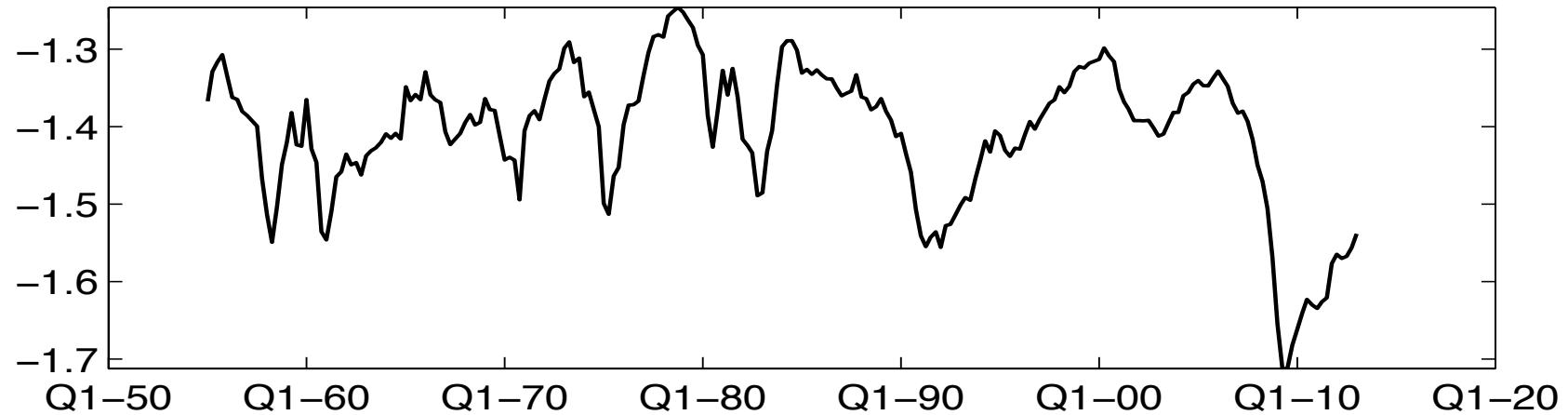


Forecasts (5 years ahead)

C-Y

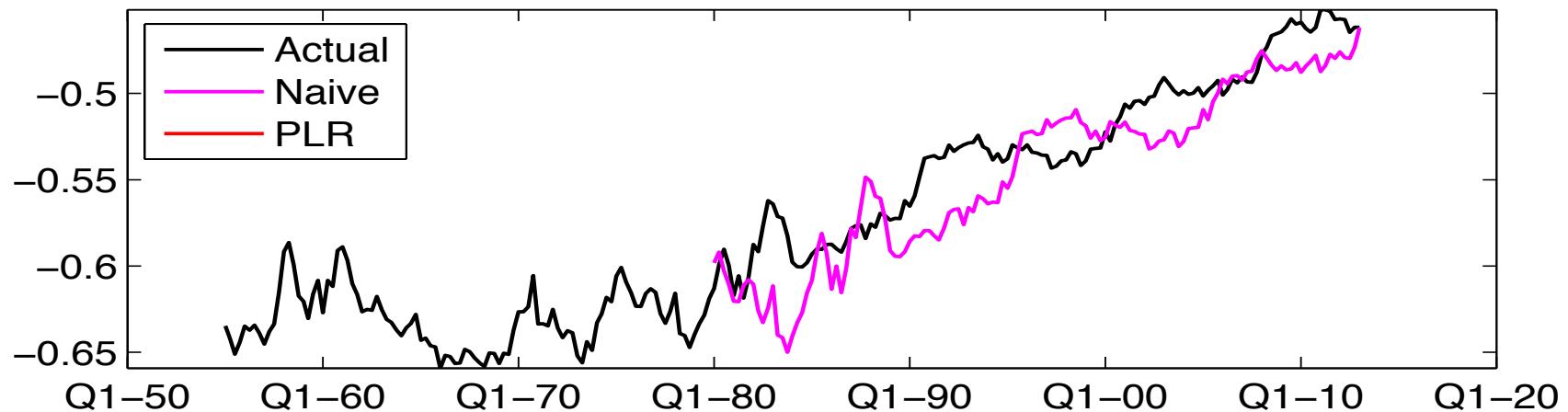


I-Y

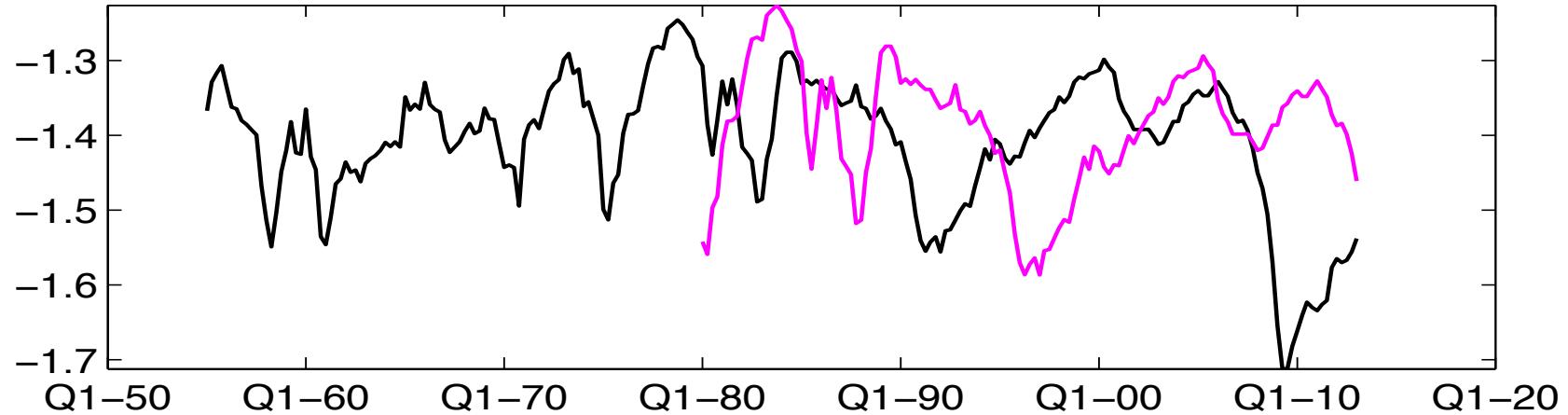


Forecasts (5 years ahead)

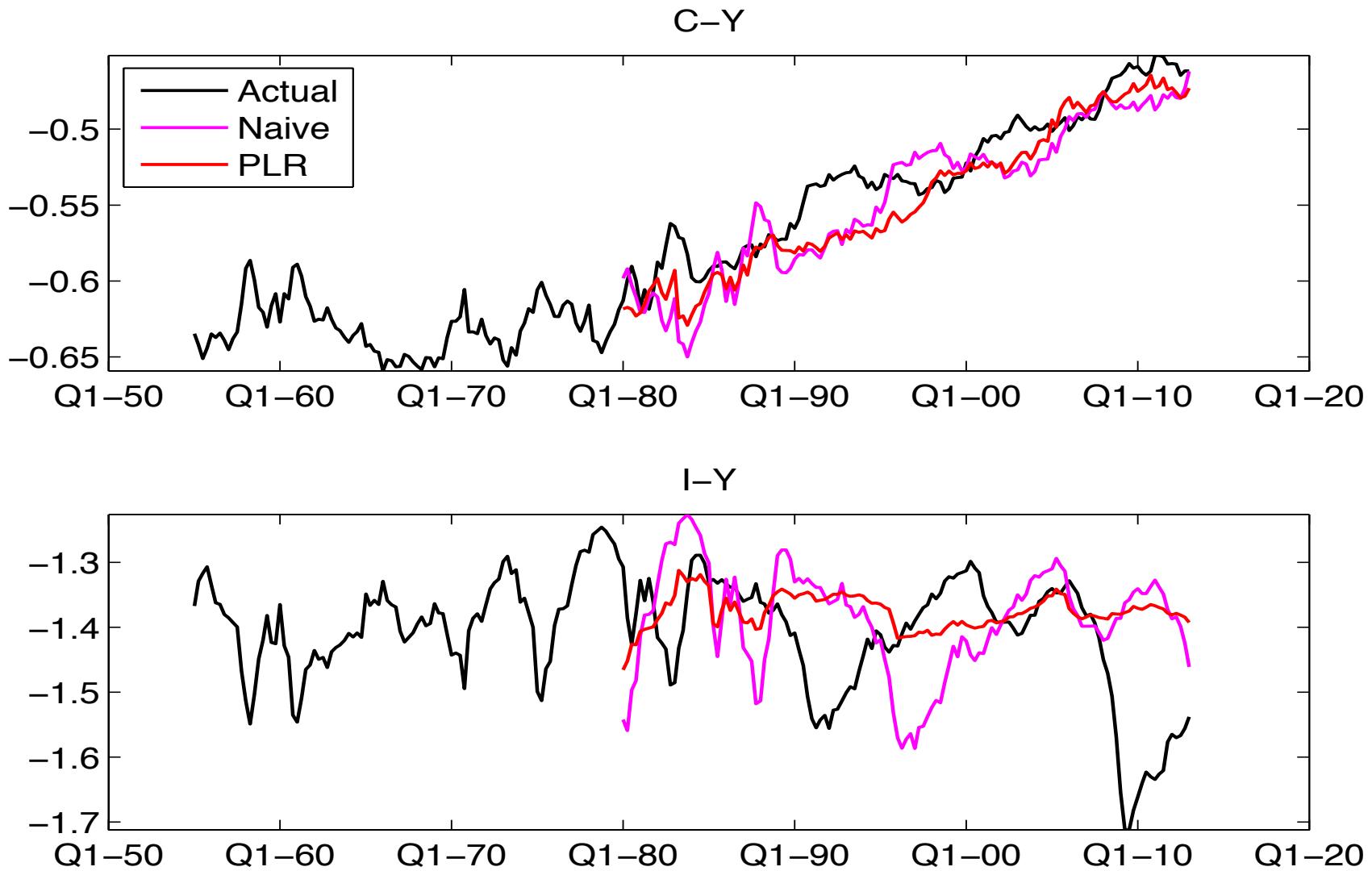
C-Y



I-Y



Forecasts (5 years ahead)



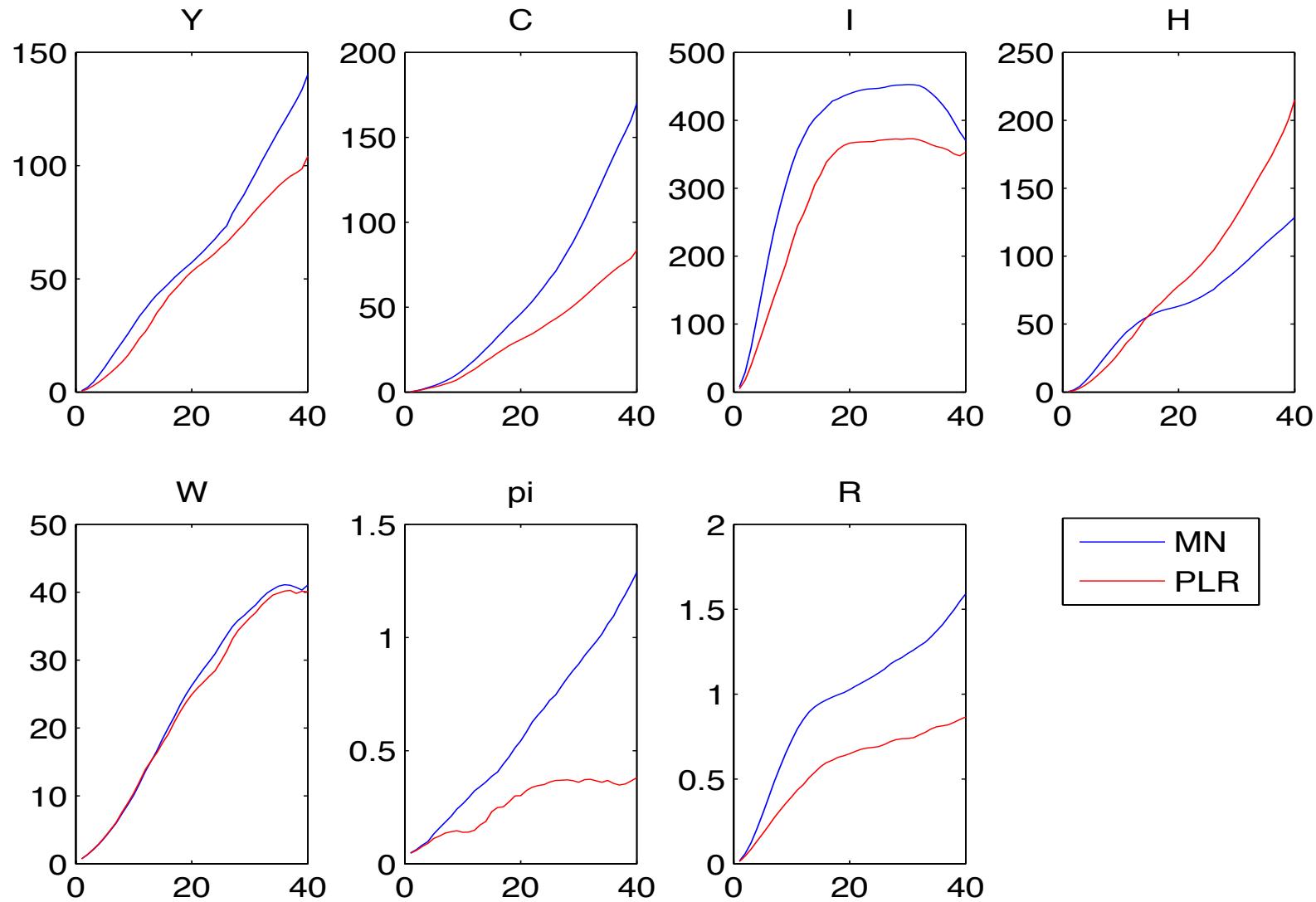
7-variable VAR

- VAR(5) with quarterly data on
- Recursive estimation starts in 1955:I
- Forecast-evaluation sample: 1985:I – 2013:I

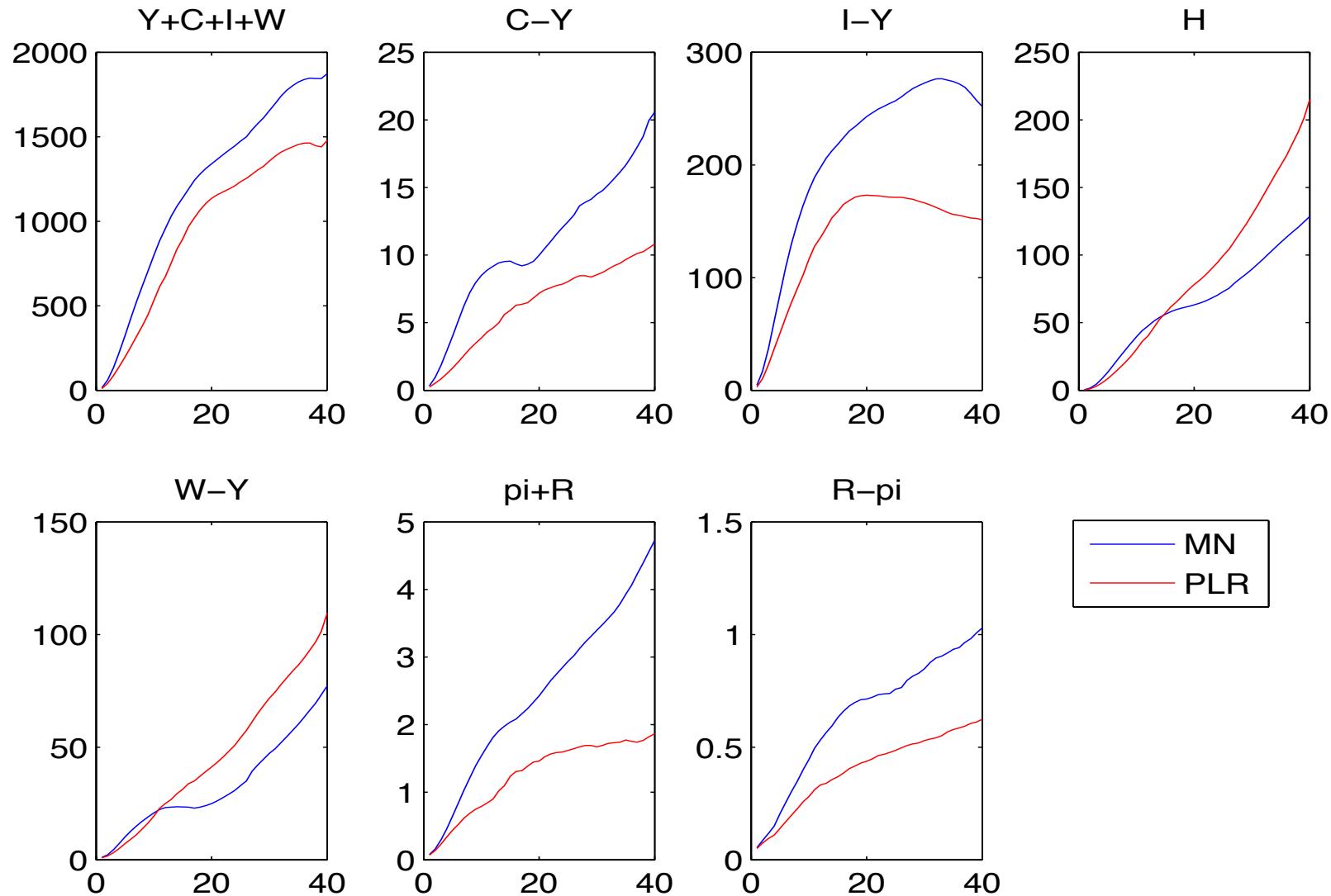
■ $H = \begin{bmatrix} x & c & i & w & h & \pi & R \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$ Interpretation of $H y$

→ Real trend
→ Nominal trend
→ Cons-GDP ratio
→ Inv-GDP ratio
→ Wage-GDP ratio
→ Hours
→ Real interest rate

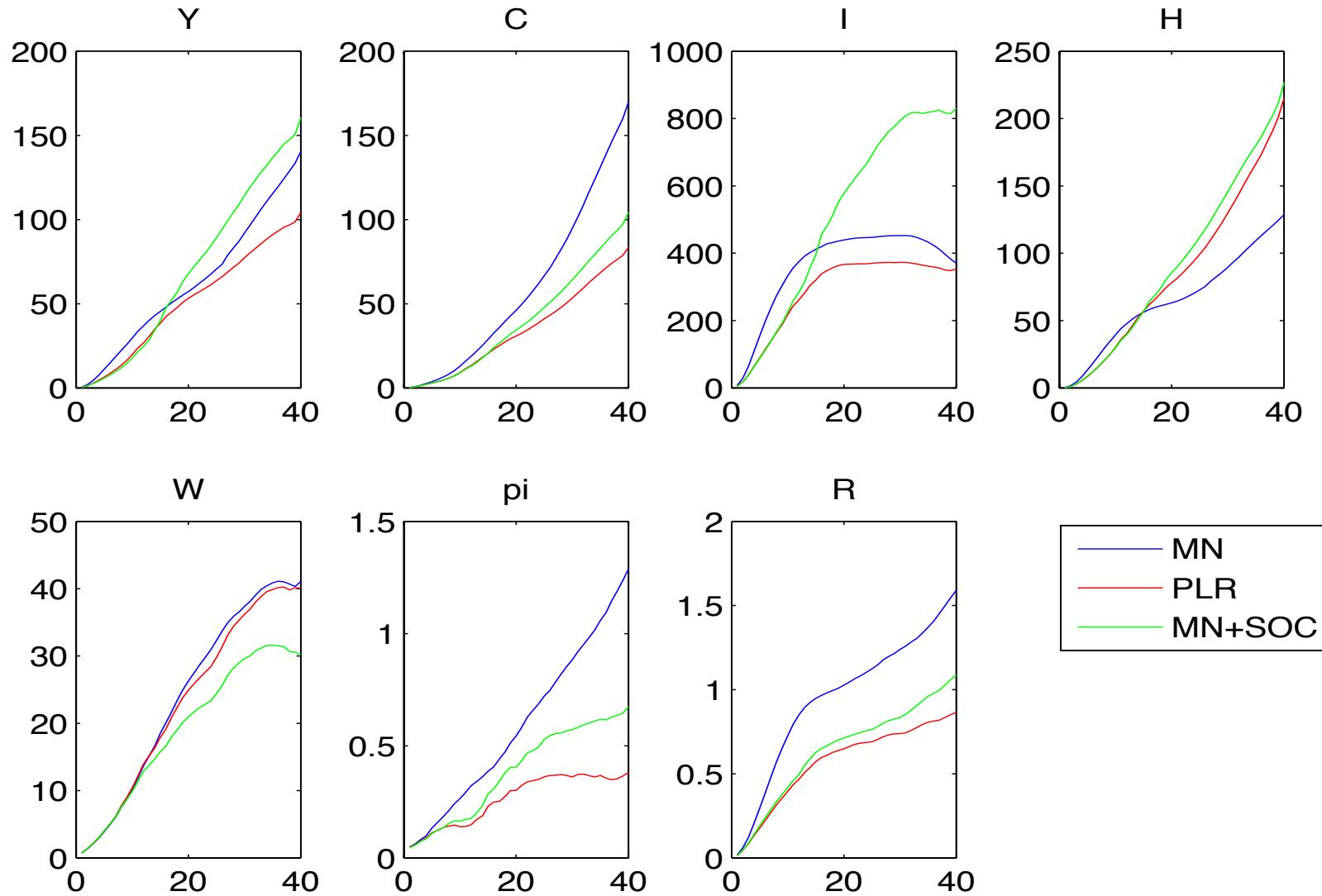
Mean Squared Forecast Errors on y (1985-2013)



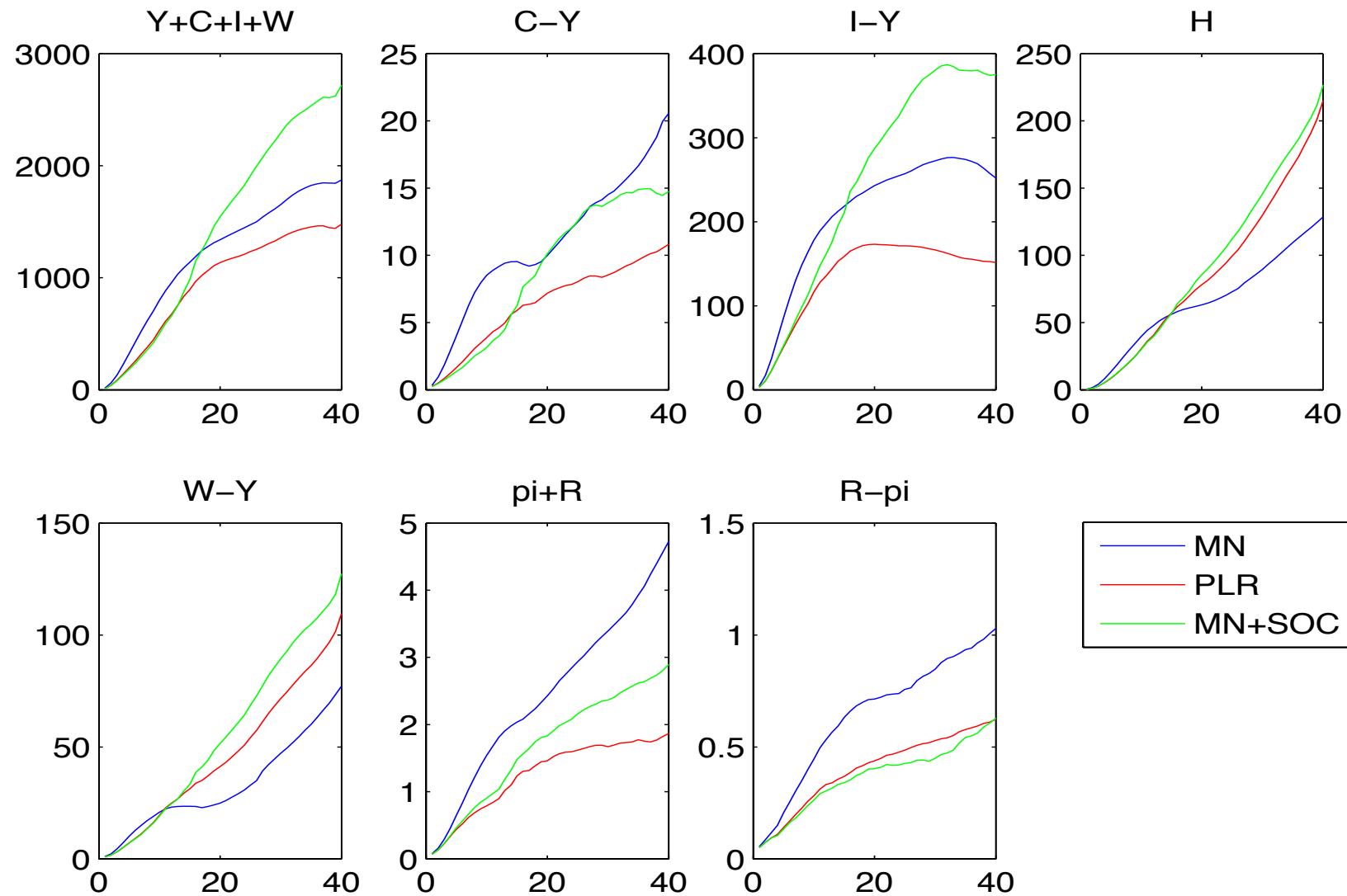
Mean Squared Forecast Errors on Hy (1985-2013)



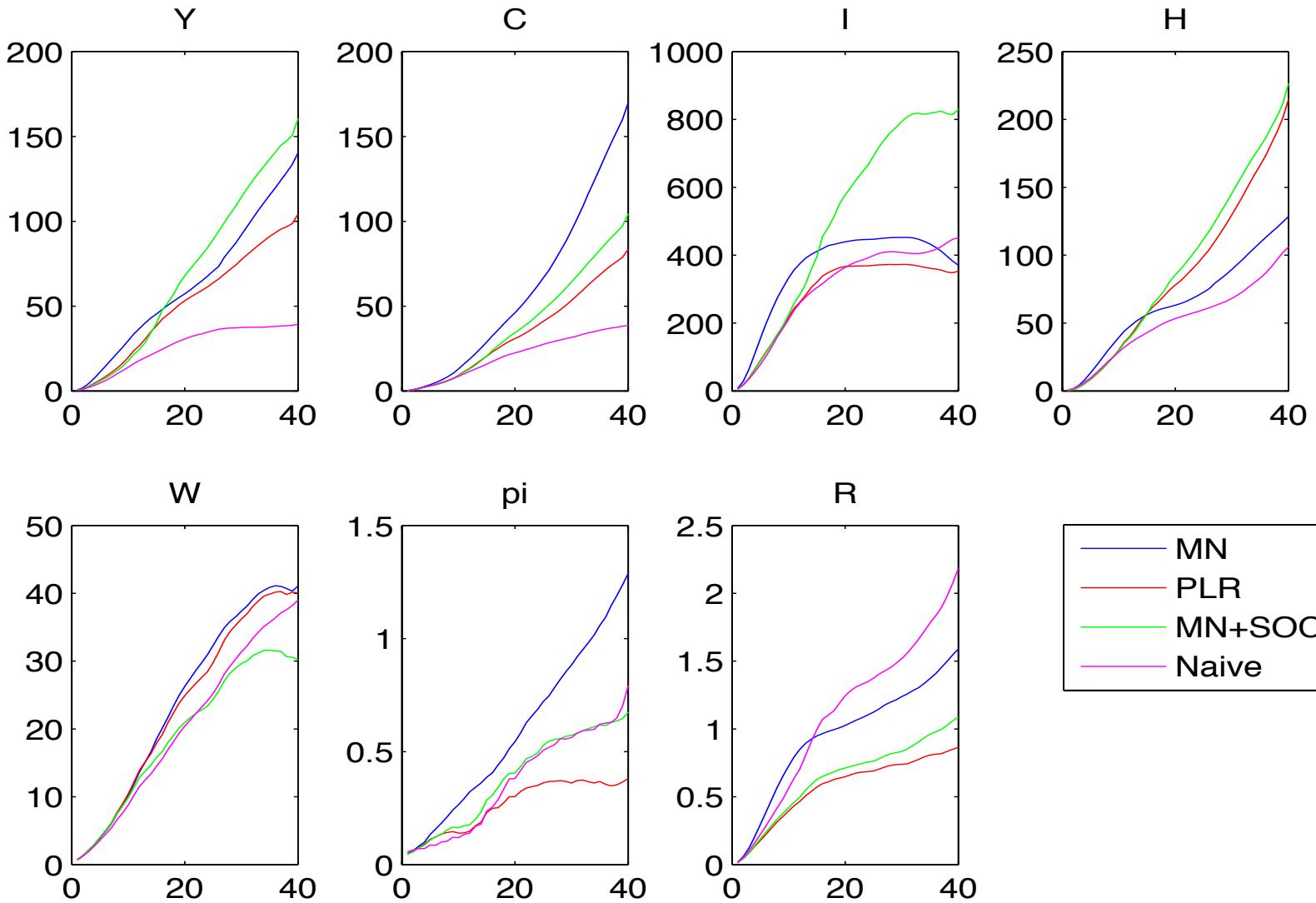
Mean Squared Forecast Errors on y (1985-2013)



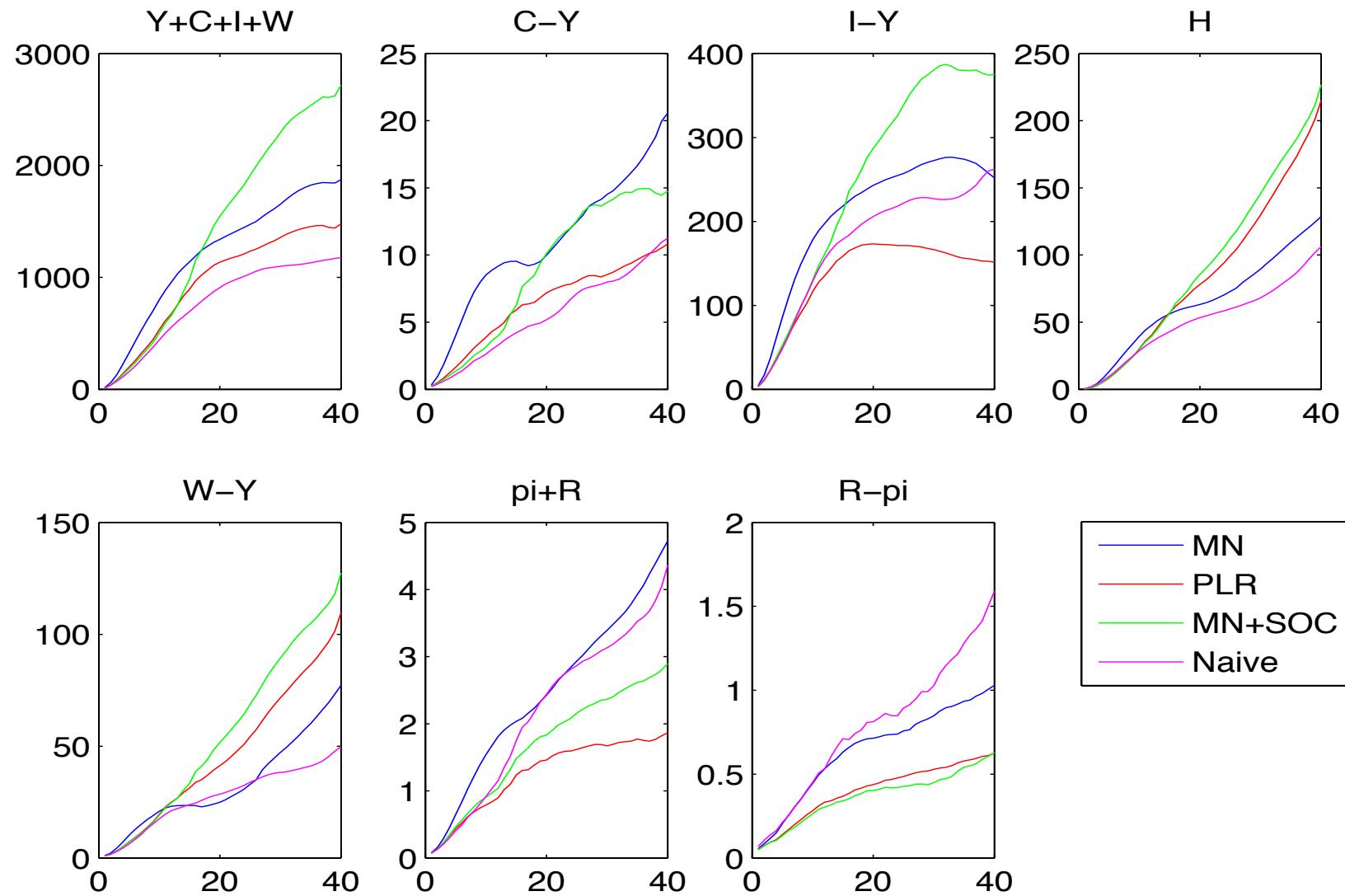
Mean Squared Forecast Errors on Hy (1985-2013)



Mean Squared Forecast Errors on y (1985-2013)



Mean Squared Forecast Errors on Hy (1985-2013)



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
- Danger of data mining in the choice of H

A prior invariant to rotations

$$\Delta y_t = c + \Lambda_1 \beta'_1 y_{t-1} + \Lambda_2 \beta'_2 y_{t-1} + \varepsilon_t$$

- Economic theory tell us what $sp(\beta)$ is likely to be, not β

A prior invariant to rotations

$$\Delta y_t = c + \Lambda_1 \beta'_1 y_{t-1} + \Lambda_2 \beta'_2 y_{t-1} + \varepsilon_t$$

- Economic theory tell us what $sp(\beta)$ is likely to be, not β
- We have also developed a prior that is invariant to rotations of β
- Important to isolate linear combinations that behave “differently”

A prior invariant to rotations

$$\Delta y_t = c + \Lambda_1 \beta'_1 y_{t-1} + \Lambda_2 \beta'_2 y_{t-1} + \varepsilon_t$$

- Economic theory tell us what $sp(\beta)$ is likely to be, not β
- We have also developed a prior that is invariant to rotations of β
- Important to isolate linear combinations that behave “differently”
- Inspired by cointegration literature, but 3 main differences
 - Fully probabilistic approach without multistep-procedure
 - Do not dogmatically shut down the loadings of non-stationary combinations
 - Shrink the loading of stationary combinations to reduce the importance of DC