Large-Scale Dynamic Predictive Regressions

Daniele Bianchi

University of Warwick
Warwick Business School

Kenichiro McAlinn

University of Chicago
Booth School of Business

2018 NBP Workshop on Forecasting, Warsaw

Set the Stage

A decision maker $\mathcal{D}$ wants to predict some quantity y based on a set of predictors $\boldsymbol{z}$.

A canonical approach is

$$y_t = \boldsymbol{\beta}' \boldsymbol{z}_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \pi(0, \nu_t)$$

Set the Stage

A decision maker $\mathcal{D}$ wants to predict some quantity y based on a set of predictors $\boldsymbol{z}$.

A canonical approach is

$$y_t = \boldsymbol{\beta}' \boldsymbol{z}_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \pi(0, \nu_t)$$

Data-rich environments:

- Availability of many predictors.
- E.g., FRED-MD for macro, firm's characteristics for finance.

Set the Stage

A decision maker $\mathcal{D}$ wants to predict some quantity y based on a set of predictors $\boldsymbol{z}$.

A canonical approach is

$$y_t = \boldsymbol{\beta}' \boldsymbol{z}_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \pi(0, \nu_t)$$

Data-rich environments:

- Availability of many predictors.
- E.g., FRED-MD for macro, firm's characteristics for finance.

Limitations of conventional approaches:

- OLS, MLE, Bayes with flat priors fail out-of-sample.
- Curse of dimensionality (over-fitting).

Set the Stage

Forecasting with many predictors:

Sparse Modeling (regularization/shrinkage)	**Dense Modeling (factor models)**

Set the Stage

Forecasting with many predictors:

Sparse Modeling
(regularization/shrinkage)

Selection of few variables with the highest explanatory power

- E.g., Lasso, Elastic-Net, SSVS (spike-and-slab).

Set the Stage

Forecasting with many predictors:

Sparse Modeling (regularization/shrinkage)

Selection of few variables with the highest explanatory power

- E.g., Lasso, Elastic-Net, SSVS (spike-and-slab).

Pros:

- Interpretability of the output.
- Parsimony.

Cons:

- Illusion of sparsity (see Giannone et al. 2018).
- Unstable over time.
- Could shrink variables that matter (for decision making).

Set the Stage

Forecasting with many predictors:

Dense Modeling
(factor models)

Few common components capture statistical features of a large set of predictors

- E.g., PCA/Factor Analysis.

Set the Stage

Forecasting with many predictors:

Pros:

- Widely used for predictions.
- Works well out-of-sample.

Cons:

- Lacks of interpretability (ex-post heuristics).
- Lacks of decision insight (un-supervised learning).

Dense Modeling (factor models)

Few common components capture statistical features of a large set of predictors

- E.g., PCA/Factor Analysis.

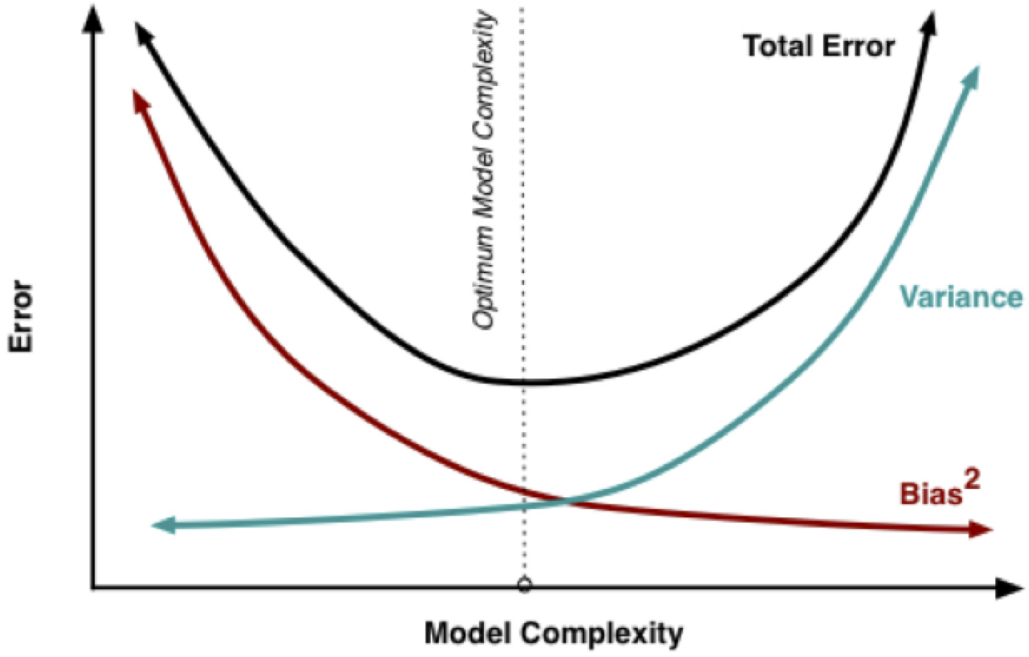
Set the Stage

Sparse Modeling
(regularization/shrinkage)

Dense Modeling
(factor modeling)

Both approaches reduce over-fitting by reducing model complexity.

Bias-variance trade-off mitigated through cross-validation or using information criteria.



Set the Stage

What we do: we propose a large-scale predictive strategy which...

- ...is based on ensembling smaller models (aggreg. of "weak learners").
- ...does not assume/impose sparsity (we retain all the data).
- ...keeps some economic interpretability (relative "weights" of predictors).

Set the Stage

What we do: we propose a large-scale predictive strategy which...

- ...is based on ensembling smaller models (aggreg. of "weak learners").
- ...does not assume/impose sparsity (we retain all the data).
- ...keeps some economic interpretability (relative "weights" of predictors).

Application(s): (1) forecasting U.S. inflation based on macro variables, and (2) forecasting industry-specific stock returns based on firms' characteristics.

Set the Stage

What we do: we propose a large-scale predictive strategy which...

- ...is based on ensembling smaller models (aggreg. of "weak learners").
- ...does not assume/impose sparsity (we retain all the data).
- ...keeps some economic interpretability (relative "weights" of predictors).

Application(s): (1) forecasting U.S. inflation based on macro variables, and (2) forecasting industry-specific stock returns based on firms' characteristics.

Results: better out-of-sample predictability compared to conventional sparse/dense predictive strategies,

- Substantial economic gains when forecasting stock returns.
- Results are rationalized in simulation.

Set the Stage

A non-exhaustive literature review:

1. Regression-based stock returns and inflation forecasting:
   (e.g., Pastor and Stambaugh (2009), Rapach et al. (2010), Van Binsbergen and Koijen (2010), Dangl and Halling (2012), Rapach et al. (2018), Stock and Watson (2002a,2002b,2006), etc.)

2. Dense modeling, regularization and sparsity:
   (e.g., Bryzgalova (2015), Clarke (2015), Chinco, Clark-Joseph, and Ye (2016), Freyberger, Neuhierl, and Weber (2017), Kozak, Nagel, and Santosh (2017), Giannone, Lenza, and Primiceri (2017), etc.)

3. Aggregation of "weak learners" and ensemble modelling:
   (e.g., West and Crosse (1992), West (1992), Gneiting, Balabdaoui, and Raftery (2007), McCulloch et al, (2010), Amisano and Geweke (2011), Casarin, Grassi, Ravazzolo and van Dijk (2018), Lee and Clyde (2004), Gruber and West (2017), McAlinn and West (2018), etc.)

Outline

1. Modeling Framework
   - Decouple-Recouple Predictive Strategy.
2. Estimation
3. Empirical Study
   - Forecasting U.S. inflation.
   - Forecasting stock returns across U.S. industries.
4. Simulation Study
5. Conclusion

Decoupling-Recoupling Predictive Strategy

$$y_t = \boldsymbol{\beta}_1' \boldsymbol{z}_{t-1,1} + ... + \boldsymbol{\beta}_j' \boldsymbol{z}_{t-1,j} + ... + \boldsymbol{\beta}_J' \boldsymbol{z}_{t-1,J} + \epsilon_t$$

Decoupling-Recoupling Predictive Strategy

$$y_t = \boldsymbol{\beta}_1' \boldsymbol{z}_{t-1,1} + \ldots + \boldsymbol{\beta}_j' \boldsymbol{z}_{t-1,j} + \ldots + \boldsymbol{\beta}_J' \boldsymbol{z}_{t-1,J} + \epsilon_t$$

$$y_t = \boldsymbol{\beta}_1' \boldsymbol{z}_{t-1,1} + \epsilon_{t,1} \qquad y_t = \boldsymbol{\beta}_j' \boldsymbol{z}_{t-1,j} + \epsilon_{t,j} \qquad y_t = \boldsymbol{\beta}_J' \boldsymbol{z}_{t-1,J} + \epsilon_{t,J}$$

Decoupling-Recoupling Predictive Strategy

$$y_t = \boldsymbol{\beta}_1' \boldsymbol{z}_{t-1,1} + \ldots + \boldsymbol{\beta}_j' \boldsymbol{z}_{t-1,j} + \ldots + \boldsymbol{\beta}_J' \boldsymbol{z}_{t-1,J} + \epsilon_t$$

$$y_t = \boldsymbol{\beta}_1' \boldsymbol{z}_{t-1,1} + \epsilon_{t,1} \qquad y_t = \boldsymbol{\beta}_j' \boldsymbol{z}_{t-1,j} + \epsilon_{t,j} \qquad y_t = \boldsymbol{\beta}_J' \boldsymbol{z}_{t-1,J} + \epsilon_{t,J}$$

$$h_1(x_1) = p(y_{t+k}|\mathcal{A}_1) \qquad h_j(x_j) = p(y_{t+k}|\mathcal{A}_j) \qquad h_J(x_J) = p(y_{t+k}|\mathcal{A}_J)$$

Decoupling-Recoupling Predictive Strategy

$$y_t = \boldsymbol{\beta}_1' \boldsymbol{z}_{t-1,1} + \ldots + \boldsymbol{\beta}_j' \boldsymbol{z}_{t-1,j} + \ldots + \boldsymbol{\beta}_J' \boldsymbol{z}_{t-1,J} + \epsilon_t$$

$$y_t = \boldsymbol{\beta}_1' \boldsymbol{z}_{t-1,1} + \epsilon_{t,1} \qquad y_t = \boldsymbol{\beta}_j' \boldsymbol{z}_{t-1,j} + \epsilon_{t,j} \qquad y_t = \boldsymbol{\beta}_J' \boldsymbol{z}_{t-1,J} + \epsilon_{t,J}$$

$$h_1(x_1) = p(y_{t+k}|\mathcal{A}_1) \qquad h_j(x_j) = p(y_{t+k}|\mathcal{A}_j) \qquad h_J(x_J) = p(y_{t+k}|\mathcal{A}_J)$$

$$p(y|\mathcal{H})$$

$$\hat{y}_{t+k}$$

Decoupling-Recoupling Predictive Strategy

$$y_t = \beta_1' \boldsymbol{z}_{t-1,1} + \ldots + \beta_j' \boldsymbol{z}_{t-1,j} + \ldots + \beta_J' \boldsymbol{z}_{t-1,J} + \epsilon_t$$

$$y_t = \beta_1' \boldsymbol{z}_{t-1,1} + \epsilon_{t,1} \qquad y_t = \beta_j' \boldsymbol{z}_{t-1,j} + \epsilon_{t,j} \qquad y_t = \beta_J' \boldsymbol{z}_{t-1,J} + \epsilon_{t,J}$$

$$h_1(x_1) = p(y_{t+k}|\mathcal{A}_1) \qquad h_j(x_j) = p(y_{t+k}|\mathcal{A}_j) \qquad h_J(x_J) = p(y_{t+k}|\mathcal{A}_J)$$

$$p(y|\mathcal{H})$$

$$\hat{y}_{t+k}$$

Issues when recoupling:

- Predictions are biased and correlated.
- You may have your own prior about y.

Supra-Bayesian

$\mathcal{D}$ has some prior information on the variable of interest $\pi(y)$

Supra-Bayesian

$\mathcal{D}$ has some prior information on the variable of interest $\pi(y)$

The forecasts

$$\mathcal{H} = \begin{cases} h_1(x_1) \\ h_2(x_2) \\ \vdots \\ h_J(x_J) \end{cases}$$

provide additional information on the quantity of interest y.

Supra-Bayesian

$\mathcal{D}$ has some prior information on the variable of interest $\pi(y)$

The forecasts

$$\mathcal{H} = \begin{cases} h_1(x_1) \\ h_2(x_2) \\ \vdots \\ h_J(x_J) \end{cases}$$

provide additional information on the quantity of interest y.

What is the posterior $p(y|\mathcal{H})$??

Bayesian Predictive Synthesis

General result:

$$p(y|\mathcal{H}) = \int \alpha(y|\boldsymbol{x})h(\boldsymbol{x})d\boldsymbol{x} \qquad \text{where} \qquad h(\boldsymbol{x}) = \prod_{j=1}^{J} h_j\left(x_j\right)$$

Lindley (1983,1985), Genest and Schervish (1985), Crosse and West (1992), West (1991,1992)

Bayesian Predictive Synthesis

General result:

$$p(y|\mathcal{H}) = \int \alpha(y|\boldsymbol{x})h(\boldsymbol{x})d\boldsymbol{x} \qquad \text{where} \qquad h(\boldsymbol{x}) = \prod_{j=1}^{J} h_j\left(x_j\right)$$

Latent Factors

Lindley (1983,1985), Genest and Schervish (1985), Crosse and West (1992), West (1991,1992)

Bayesian Predictive Synthesis

General result:

$$p(y|\mathcal{H}) = \int \alpha(y|\boldsymbol{x})h(\boldsymbol{x})d\boldsymbol{x} \qquad \text{where} \qquad h(\boldsymbol{x}) = \prod_{j=1}^{J} h_j(x_j)$$

Synthesis Function
(Conditional Specification)

Lindley (1983,1985), Genest and Schervish (1985),
Crosse and West (1992), West (1991,1992)

Bayesian Predictive Synthesis

General result:

$$p(y|\mathcal{H}) = \int \alpha(y|\boldsymbol{x}) h(\boldsymbol{x}) d\boldsymbol{x} \qquad \text{where} \qquad h(\boldsymbol{x}) = \prod_{j=1}^{J} h_j(x_j)$$

! Necessary condition !

$$p(y) = \int \alpha(y|\boldsymbol{x}) m(\boldsymbol{x}) d\boldsymbol{x}$$

$$m(\boldsymbol{x}) = E\left[\prod_{j=1}^{J} h_j(x_j)\right]$$

Lindley (1983,1985), Genest and Schervish (1985),
Crosse and West (1992), West (1991,1992)

Bayesian Predictive Synthesis

General result:

$$p(y|\mathcal{H}) = \int \alpha(y|\boldsymbol{x}) h(\boldsymbol{x}) d\boldsymbol{x} \qquad \text{where} \qquad h(\boldsymbol{x}) = \prod_{j=1}^{J} h_j(x_j)$$

! Necessary condition !

N.B: The functional form of $\alpha(y|\boldsymbol{x})$ is left unspecified by the theory.

$$p(y) = \int \alpha(y|\boldsymbol{x}) m(\boldsymbol{x}) d\boldsymbol{x}$$

$$m(\boldsymbol{x}) = E\left[\prod_{j=1}^{J} h_j(x_j)\right]$$

Lindley (1983,1985), Genest and Schervish (1985), Crosse and West (1992), West (1991,1992)

Dynamic Synthesis Function

We follow McAlinn and West (2018) and McAlinn et al. (2018) and assume

$$p(y_t|\boldsymbol{\Phi}_t, \mathcal{H}_t) = \int \alpha(y_t|\boldsymbol{x}_t, \boldsymbol{\Phi}_t) h_t(\boldsymbol{x}_t) d\boldsymbol{x}_t$$

$\boldsymbol{x}_t = (x_1, \ldots, x_J)'$ vector of latent states at time t

Dynamic Synthesis Function

We follow McAlinn and West (2018) and McAlinn et al. (2018) and assume

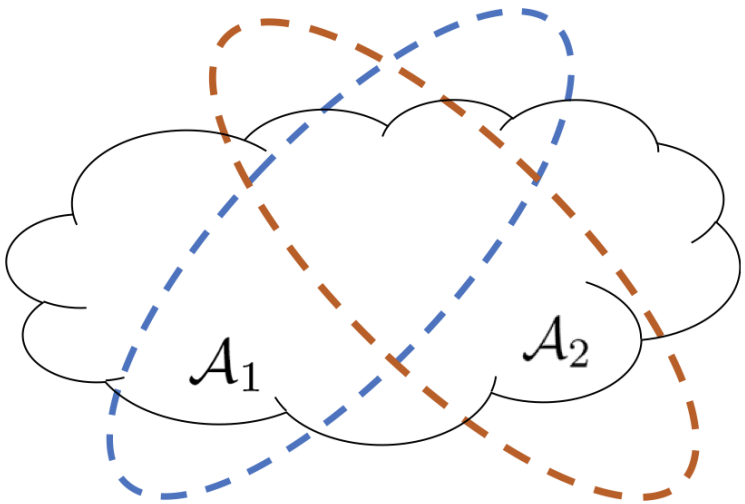
$$p(y_t|\boldsymbol{\Phi}_t, \mathcal{H}_t) = \int \alpha(y_t|\boldsymbol{x}_t, \boldsymbol{\Phi}_t) h_t(\boldsymbol{x}_t) d\boldsymbol{x}_t$$

Dynamic Latent Factor Model

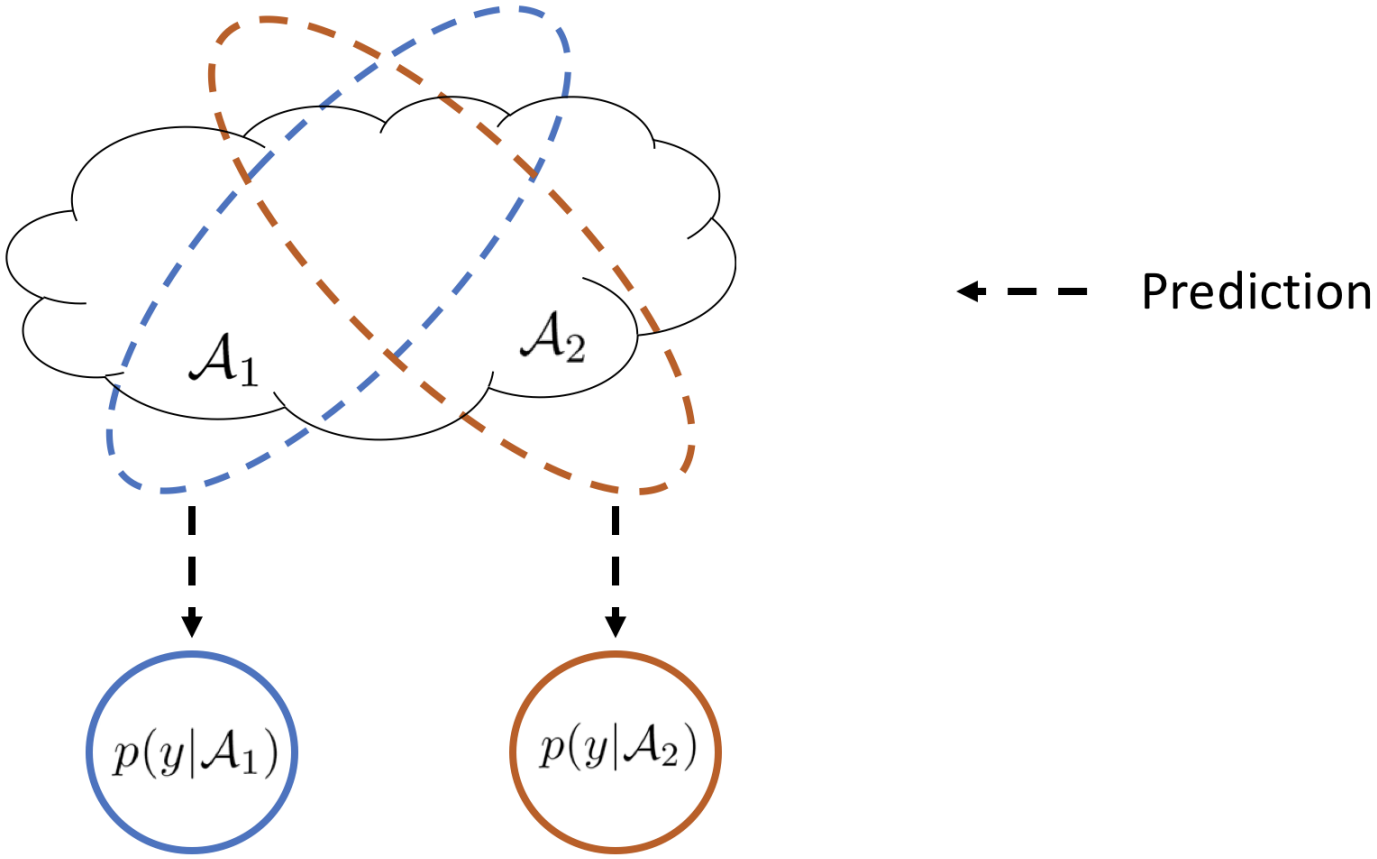
$\boldsymbol{x}_t = (x_1, \ldots, x_J)'$ vector of latent states at time t

$$y_t = \boldsymbol{F}_t' \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N(0, v_t),$$
$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(0, v_t \boldsymbol{W}_t),$$
$$\boldsymbol{\Phi}_t = (\boldsymbol{\theta}_t, v_t)$$
$$\boldsymbol{F}_t = (1, \boldsymbol{x}_t')' \quad \text{and} \quad \boldsymbol{\theta}_t = (\theta_{t0}, \theta_{t1}, ..., \theta_{tJ})'$$

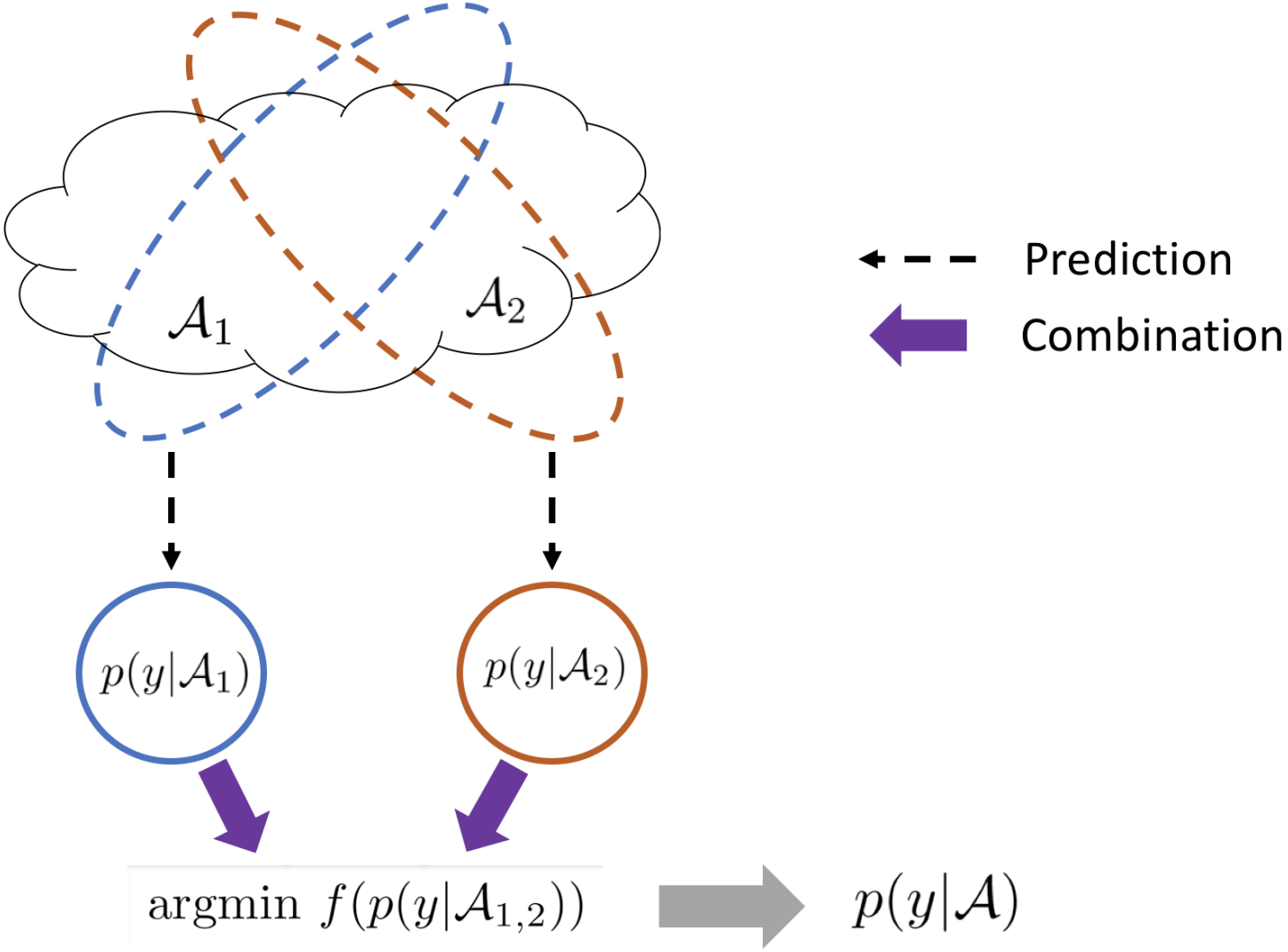
Digression: Yet Again Model Combination?



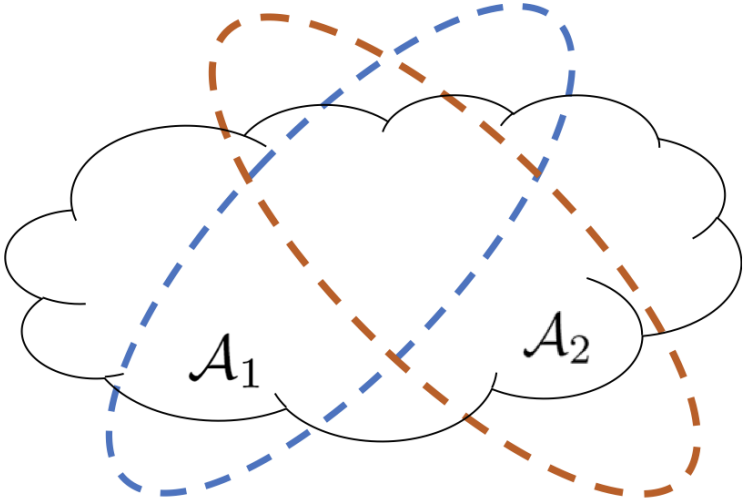
Digression: Yet Again Model Combination?



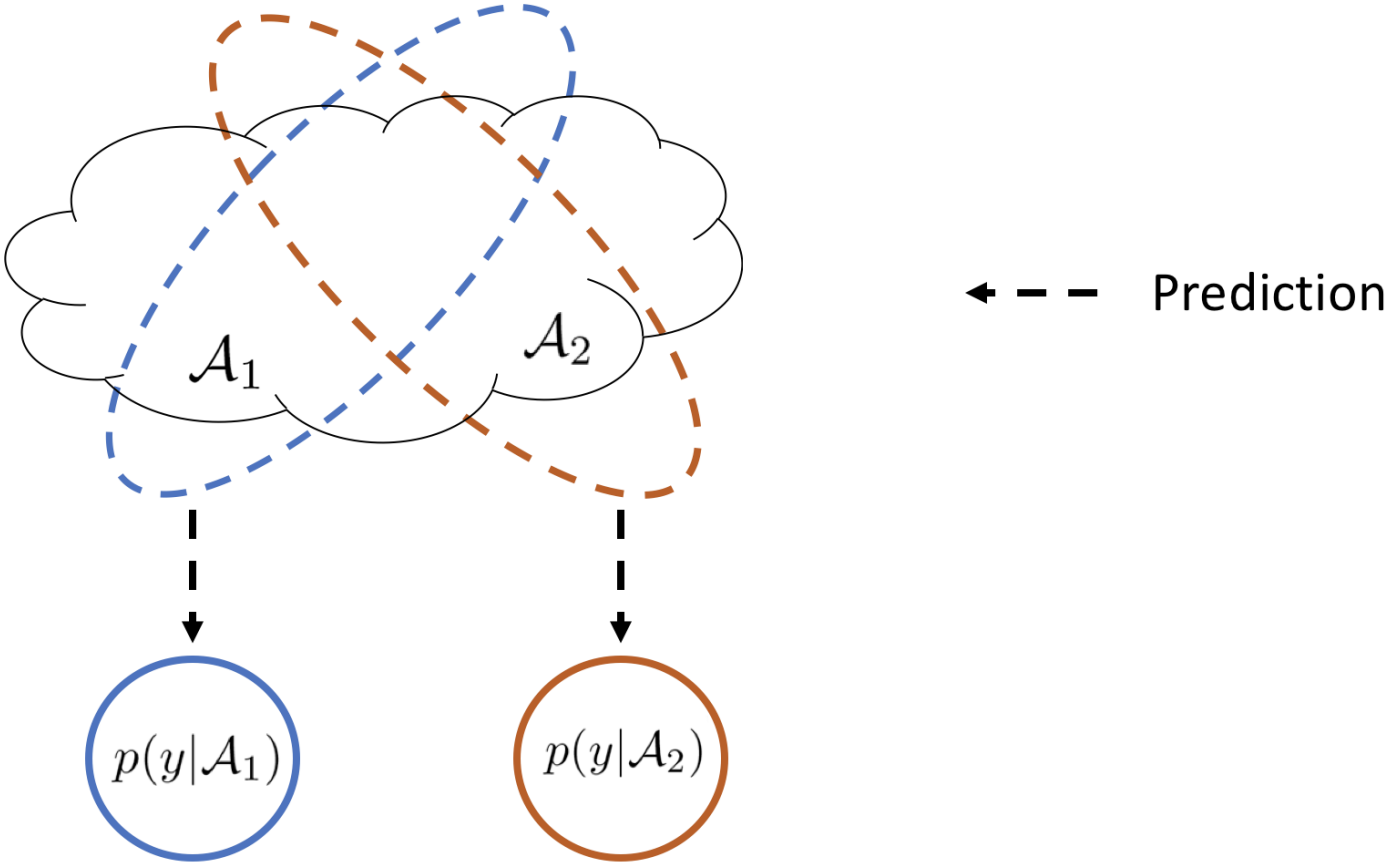
Digression: Yet Again Model Combination?



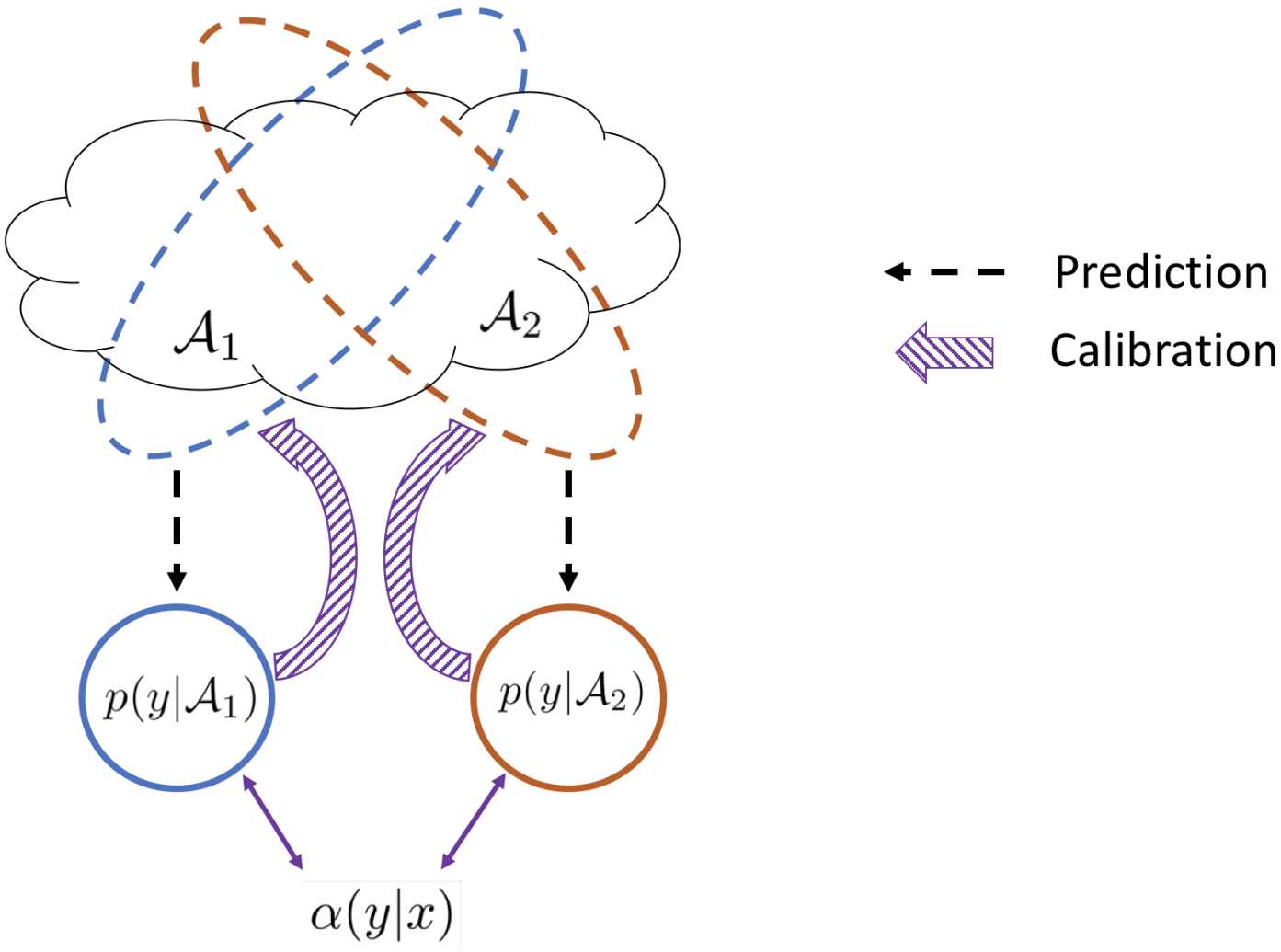
Digression: Yet Again Model Combination?



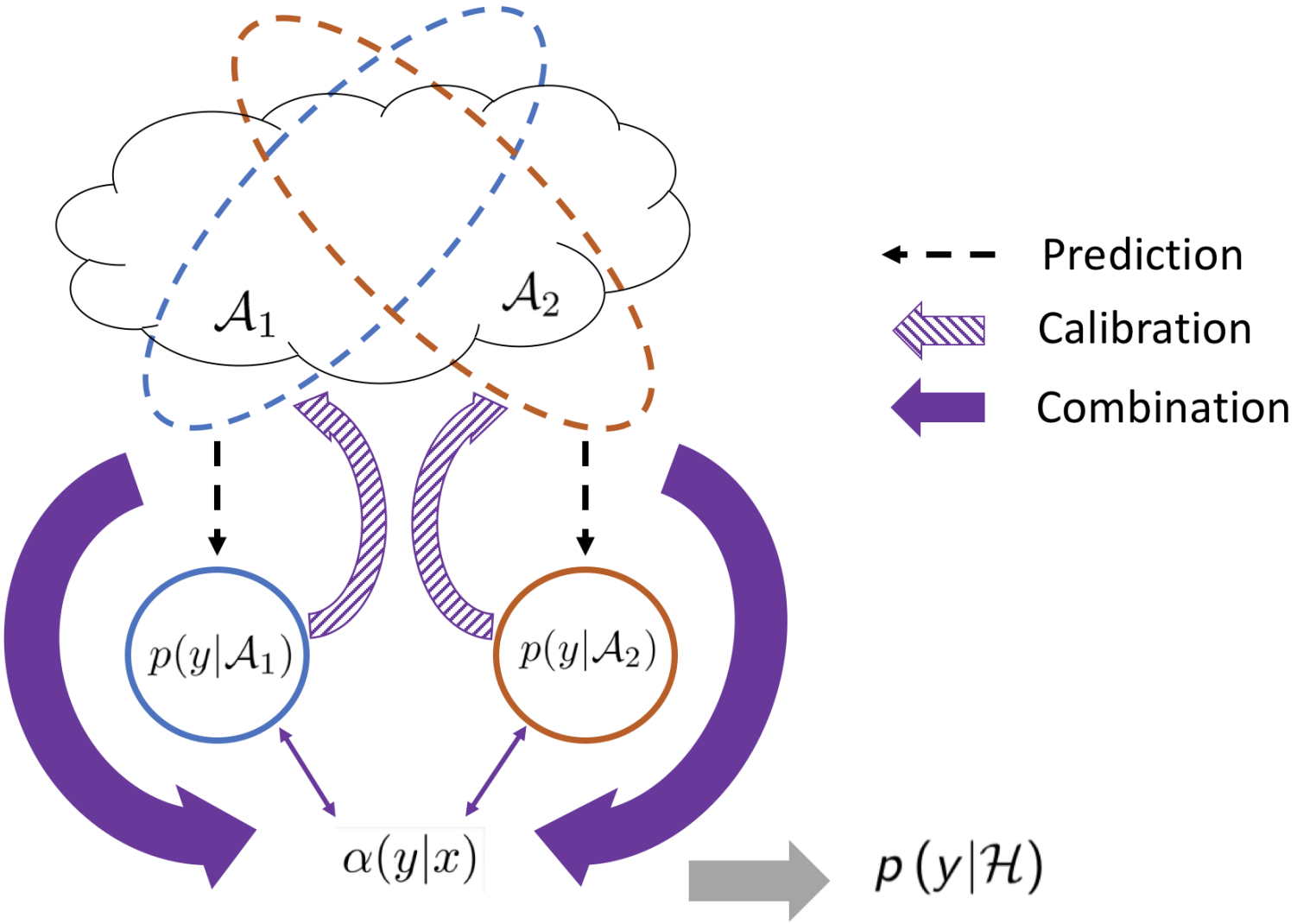
Digression: Yet Again Model Combination?



Digression: Yet Again Model Combination?



Digression: Yet Again Model Combination?



Estimation Strategy

Two-component block Gibbs sampler:

Block 1:
(bias and interdependencies)

Conditional on the "latent states" $\boldsymbol{x}_t$, the parameters $\boldsymbol{\theta}_t$, v_t and $\boldsymbol{W}_t$ are sampled from a Forward Filtering Backward Sampler (FFBS) algorithm.

Estimation Strategy

Two-component block Gibbs sampler:

Block 1:
(bias and interdependencies)

Conditional on the "latent states" $\boldsymbol{x}_t$, the parameters $\boldsymbol{\theta}_t$, v_t and $\boldsymbol{W}_t$ are sampled from a Forward Filtering Backward Sampler (FFBS) algorithm.

Block 2:
("latent states")

Conditional on $\boldsymbol{\theta}_t$, v_t and $\boldsymbol{W}_t$, the "latent states" $\boldsymbol{x}_t$ are sampled from

$$\propto N\left(y_t | \mathbf{F}_t' \boldsymbol{\theta}_t, v_t\right) \prod_{j=1:J} h_{tj}(x_{tj}) dx_{tj}$$

Estimation Strategy

Two-component block Gibbs sampler:

Block 1:
(bias and interdependencies)

Conditional on the "latent states" $\boldsymbol{x}_t$, the parameters $\boldsymbol{\theta}_t$, v_t and $\boldsymbol{W}_t$ are sampled from a Forward Filtering Backward Sampler (FFBS) algorithm.

Block 2:
("latent states")

Conditional on $\boldsymbol{\theta}_t$, v_t and $\boldsymbol{W}_t$, the "latent states" $\boldsymbol{x}_t$ are sampled from

$$\propto N\left(y_t | \mathbf{F}_t' \boldsymbol{\theta}_t, v_t\right) \prod_{j=1:J} h_{tj}(x_{tj}) dx_{tj}$$

N.B: With $h_{tj}(x_{tj})$ Student-t density, we use a scale mixture of normals which allows for conjugate updating.

Empirical Study (Setting)

Two (regression-based) forecasting exercises:

- Forecasting U.S. inflation based on macroeconomic variables.
- Forecasting stock returns based on firms' characteristics.

Empirical Study (Setting)

Two (regression-based) forecasting exercises:

- Forecasting U.S. inflation based on macroeconomic variables.
- Forecasting stock returns based on firms' characteristics.

Each $h_j(x_j) = p(y|\mathcal{A}_j)$ is generated by a TVP + SV for $j = 1, \ldots, J$.

$$y_t = \boldsymbol{\beta}_{tj}' \boldsymbol{z}_{tj} + \epsilon_{tj}, \quad \epsilon_{tj} \sim N(0, \nu_{tj}),$$
$$\boldsymbol{\beta}_{tj} = \boldsymbol{\beta}_{t-1,j} + \boldsymbol{u}_{tj}, \quad \boldsymbol{u}_{tj} \sim N(0, \nu_{tj} \boldsymbol{U}_{tj}),$$

Empirical Study (Setting)

Two (regression-based) forecasting exercises:

- Forecasting U.S. inflation based on macroeconomic variables.
- Forecasting stock returns based on firms' characteristics.

Each $h_j(x_j) = p(y|\mathcal{A}_j)$ is generated by a TVP + SV for $j = 1, \ldots, J$.

$$y_t = \boldsymbol{\beta}_{tj}' \boldsymbol{z}_{tj} + \epsilon_{tj}, \quad \epsilon_{tj} \sim N(0, \nu_{tj}),$$
$$\boldsymbol{\beta}_{tj} = \boldsymbol{\beta}_{t-1,j} + \boldsymbol{u}_{tj}, \quad \boldsymbol{u}_{tj} \sim N(0, \nu_{tj} \boldsymbol{U}_{tj}),$$

Comparison based on the cumulative Log Predictive Density Ratios (LPDR),

$$\mathrm{LPDR}_t(k) = \sum_{i=1}^{t} \log\{p(y_{i+k}|y_{1:i}, \mathcal{M}_s)/p(y_{i+k}|y_{1:i}, \mathcal{M}_0)\},$$

where $\mathcal{M}_0$ is our strategy and $\mathcal{M}_s$ denotes the competitor.

Empirical Study (Setting)

Two (regression-based) forecasting exercises:

- Forecasting U.S. inflation based on macroeconomic variables.
- Forecasting stock returns based on firms' characteristics.

N.B: The J groups are selected based on their economic meaning.

Each $h_j(x_j) = p(y|\mathcal{A}_j)$ is generated by a TVP + SV for $j = 1, \ldots, J$.

$$y_t = \boldsymbol{\beta}'_{tj}\boldsymbol{z}_{tj} + \epsilon_{tj}, \quad \epsilon_{tj} \sim N(0, \nu_{tj}),$$
$$\boldsymbol{\beta}_{tj} = \boldsymbol{\beta}_{t-1,j} + \boldsymbol{u}_{tj}, \quad \boldsymbol{u}_{tj} \sim N(0, \nu_{tj}\boldsymbol{U}_{tj}),$$

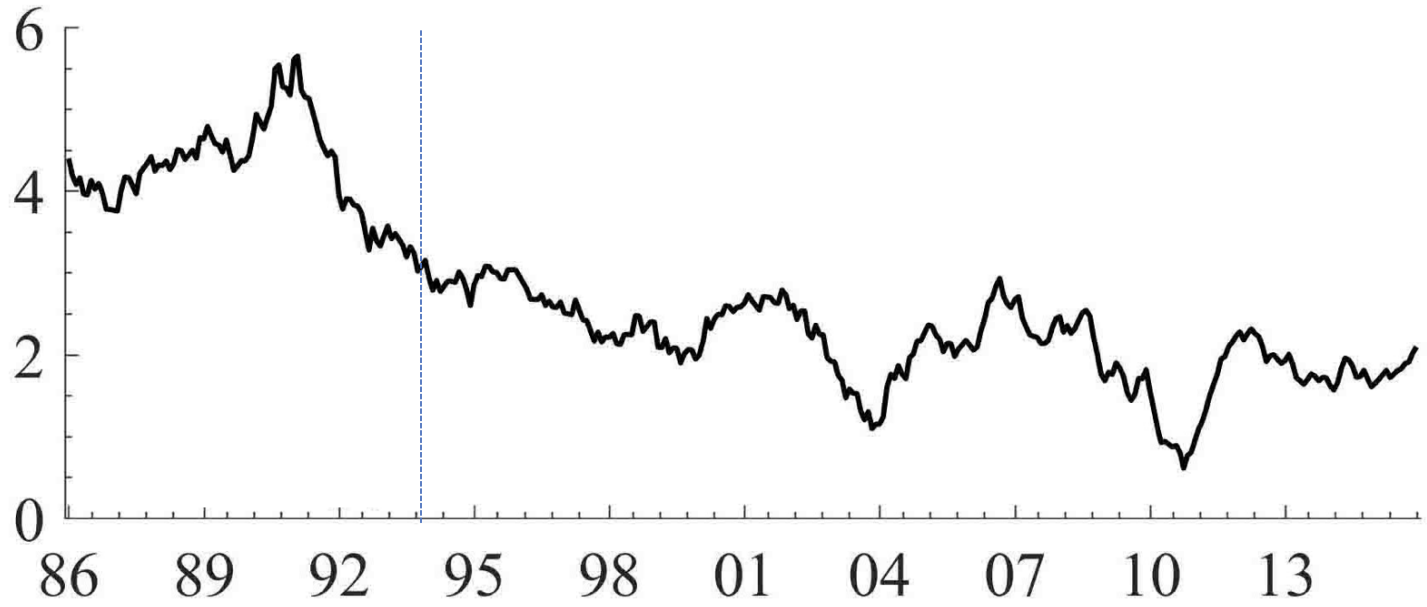
Comparison based on the cumulative Log Predictive Density Ratios (LPDR),

$$\mathrm{LPDR}_t(k) = \sum_{i=1}^{t} \log\{p(y_{i+k}|y_{1:i}, \mathcal{M}_s)/p(y_{i+k}|y_{1:i}, \mathcal{M}_0)\},$$

where $\mathcal{M}_0$ is our strategy and $\mathcal{M}_s$ denotes the competitor.

Empirical Study (Setting)

Timeline of the empirical application



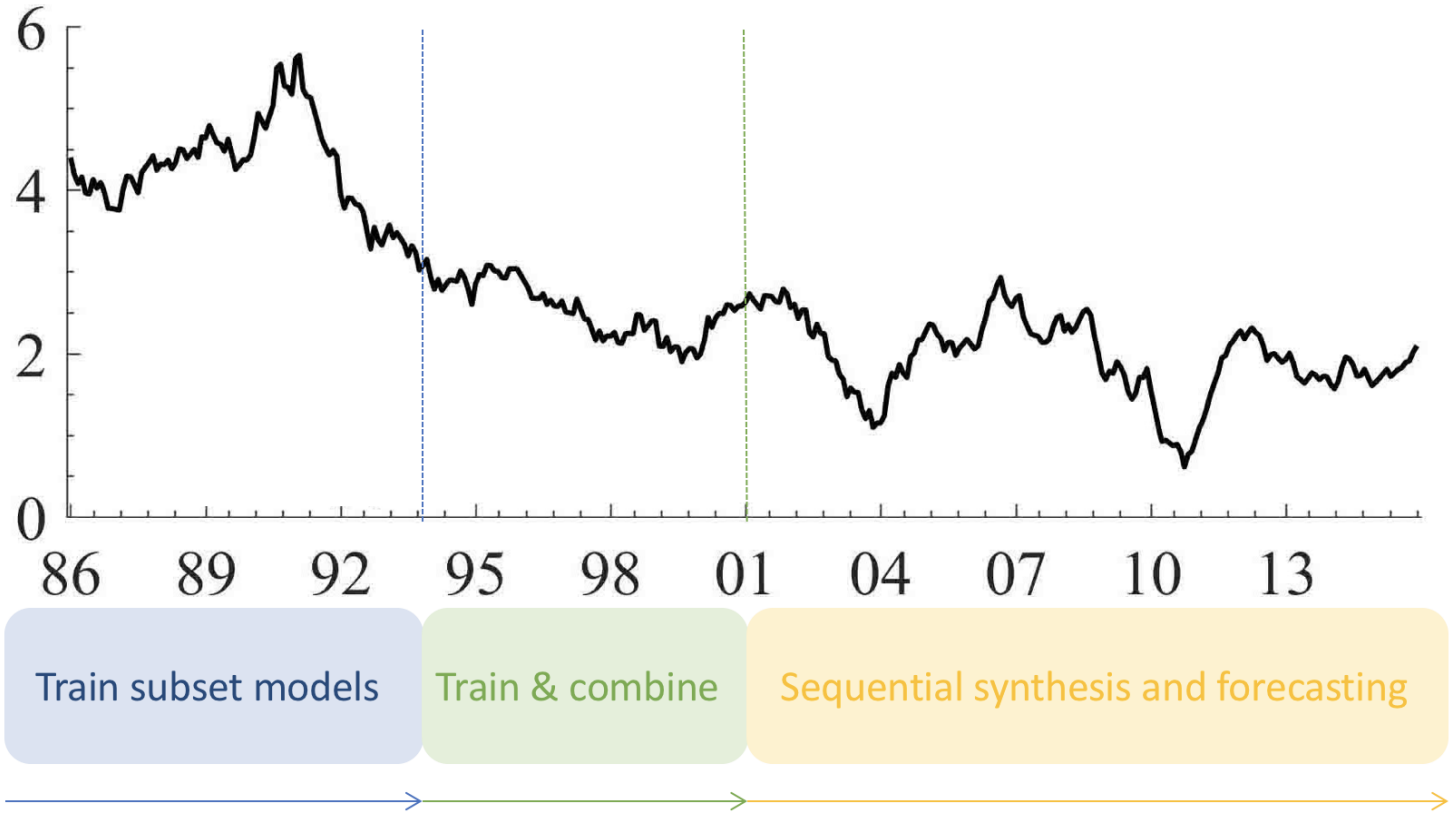
Train subset models

Empirical Study (Setting)

Timeline of the empirical application

Empirical Study (Setting)

Timeline of the empirical application



Empirical Study 1: Forecasting U.S. Inflation

Forecasting U.S. inflation using a large panel of macroeconomic variables (FRED-MD, $p = 128$)

- 1,3, and 12 months ahead

Empirical Study 1: Forecasting U.S. Inflation

Forecasting U.S. inflation using a large panel of macroeconomic variables (FRED-MD, $p = 128$)

- 1,3, and 12 months ahead

$J = 8$ groups of predictors

- Output and income
- Labor market
- Consumption
- Inventories
- Money and credit
- Interest rates and FX
- Prices
- Stock market

Empirical Study 1: Forecasting U.S. Inflation

Forecasting U.S. inflation using a large panel of macroeconomic variables (FRED-MD, $p = 128$)

- 1,3, and 12 months ahead

Compare against:

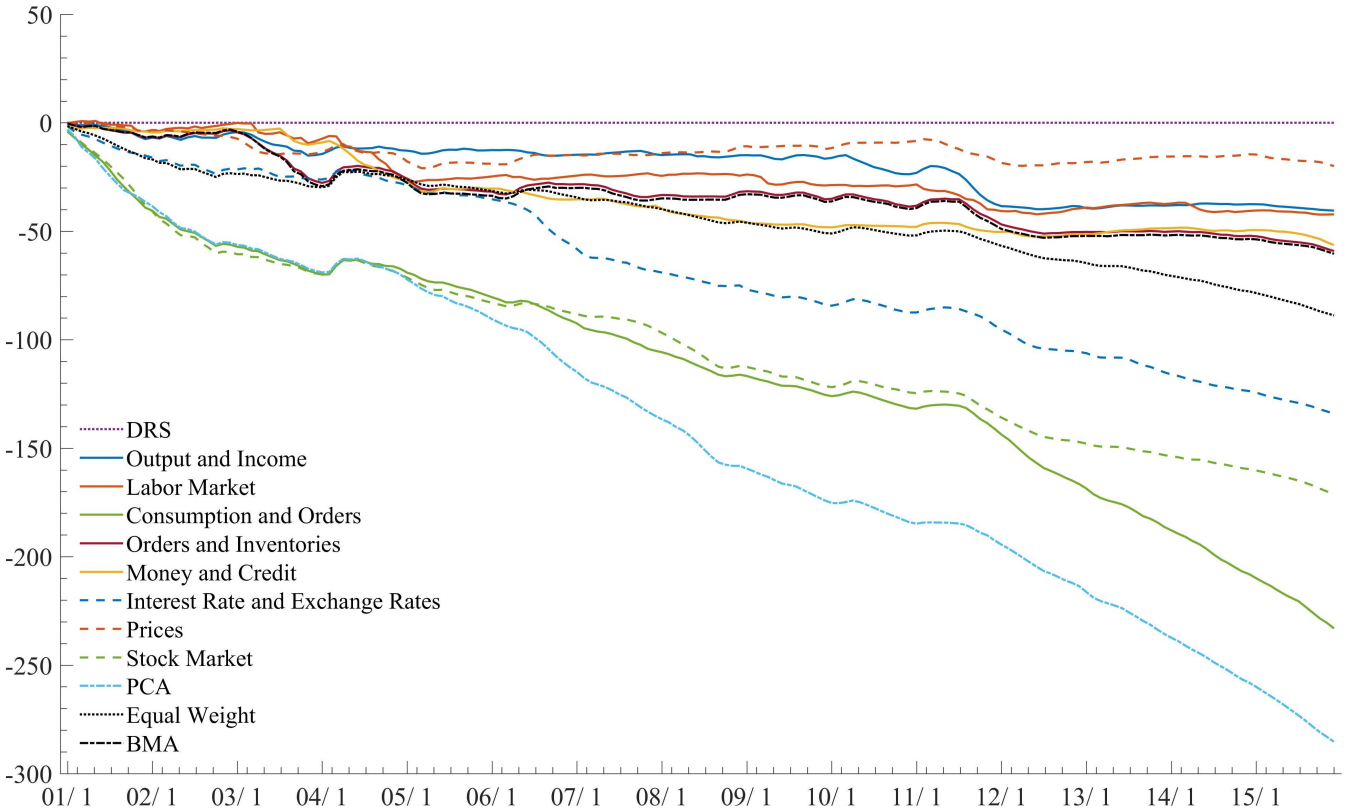
- Sequential BMA
- TVP factor model (k chosen by IC)
- Lasso (leave-one-out cross-validation)
- Linear pooling (equal-weight)
- Sub-group predictions

$J = 8$ groups of predictors

- Output and income
- Labor market
- Consumption
- Inventories
- Money and credit
- Interest rates and FX
- Prices
- Stock market

Empirical Study 1: Forecasting U.S. Inflation

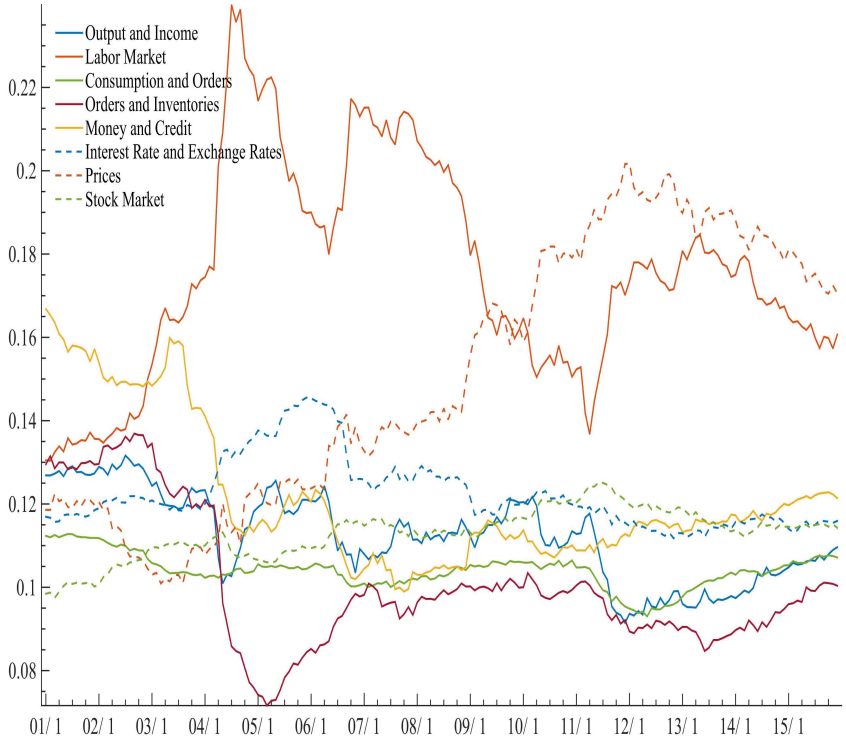
Figure: Out-of-sample log predictive density ratio



This figure shows the dynamics of the out-of-sample Log Predictive Density Ratio (LPDR) obtained for each competing specification. The sample period is 01:2001-12:2015, monthly. The objective function is the **one-step** ahead density forecast of annual inflation.

Empirical Study 1: Forecasting U.S. Inflation

Figure: Posterior means of **rescaled** interdependencies.

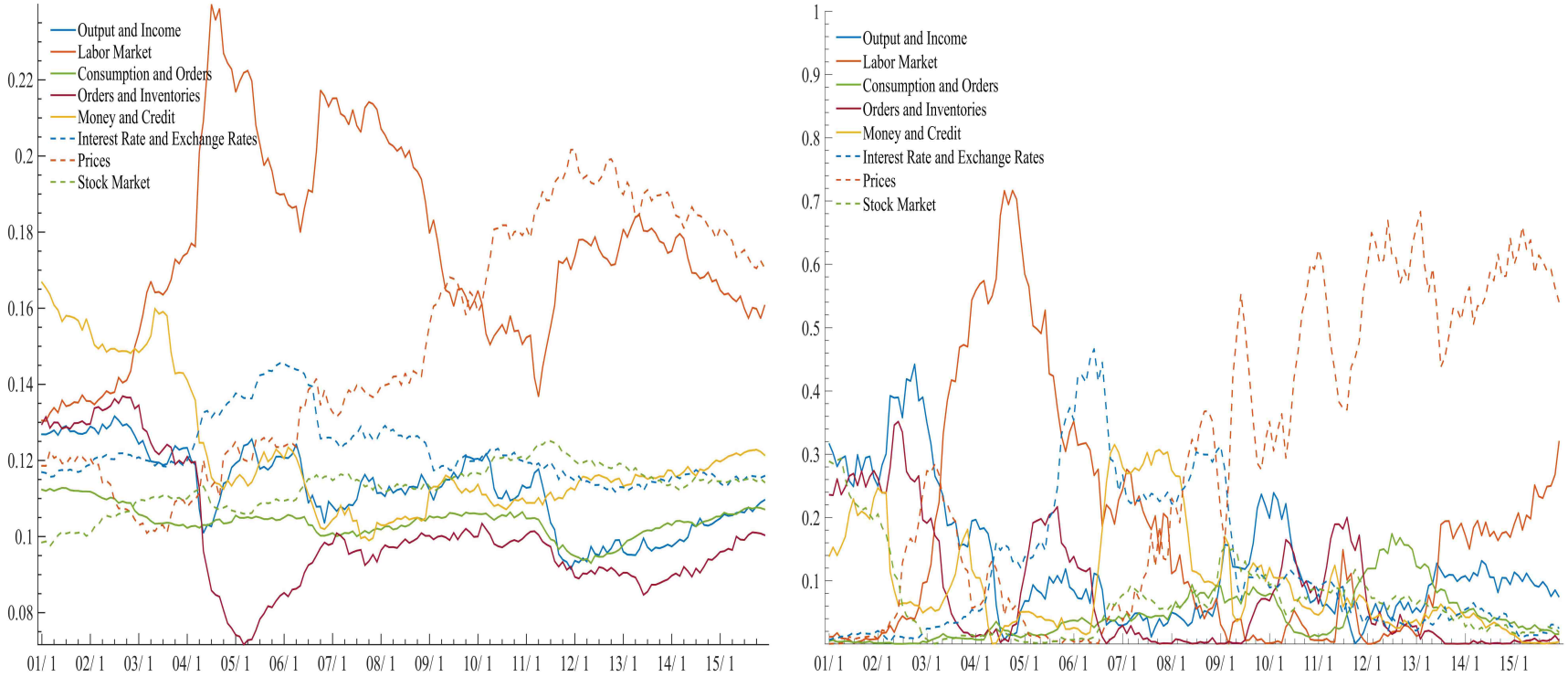


(a) 1-month ahead

This figure shows the latent interdependencies across groups of predictive densities used in the recoupling step. Left panel the posterior mean estimates for the one-step ahead forecast and right panel for the three-step ahead forecast. The parameters are rescaled between 0 and 1 via logistic transformation.

Empirical Study 1: Forecasting U.S. Inflation

Figure: Posterior means of **rescaled** interdependencies.

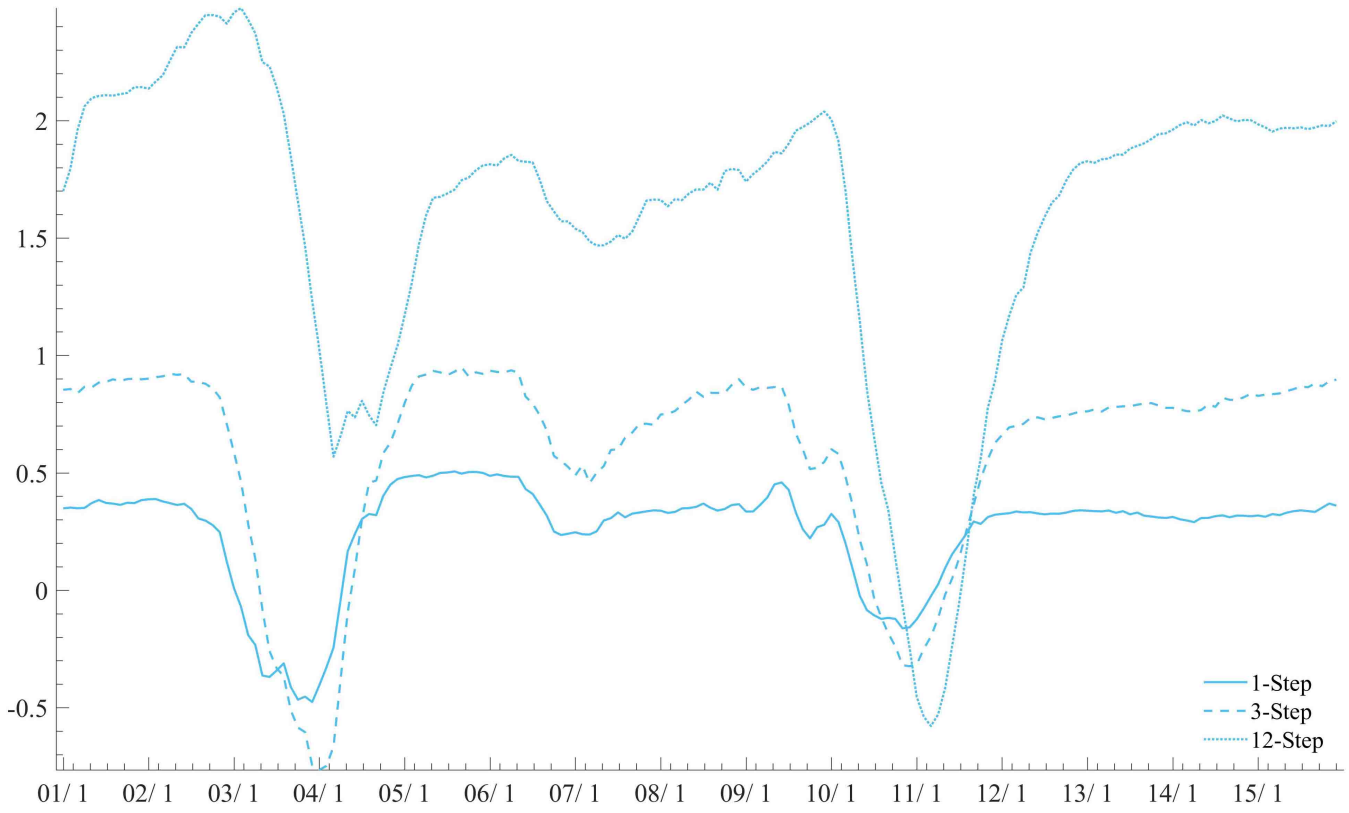


(a) 1-month ahead (b) 3-month ahead

This figure shows the latent interdependencies across groups of predictive densities used in the recoupling step. Left panel the posterior mean estimates for the one-step ahead forecast and right panel for the three-step ahead forecast. The parameters are rescaled between 0 and 1 via logistic transformation.

Empirical Study 1: Forecasting U.S. Inflation

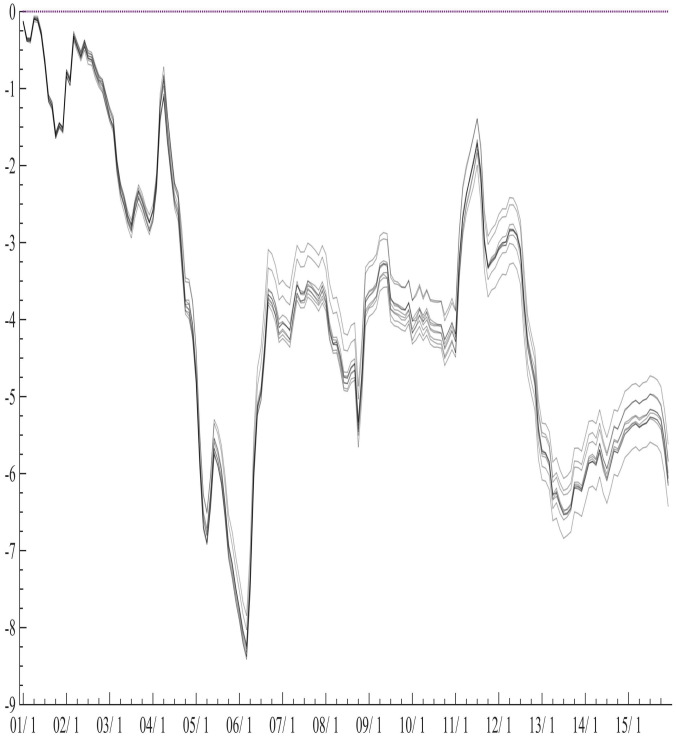
Figure: Out-of-Sample Aggregate Bias



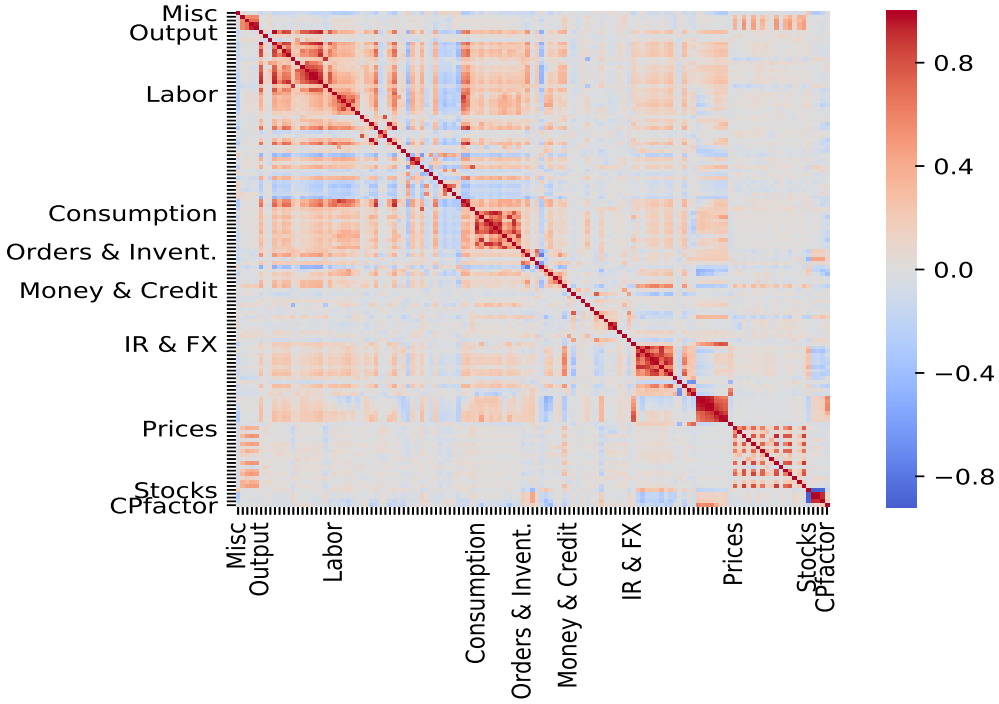
This figure shows the dynamics of the out-of-sample predictive bias obtained as the time-varying intercept from the recoupling step of the decouple-recouple dynamic predictive strategy. The sample period is 01:2001-12:2015, monthly. The objective function is the one-, three-, and twelve-step ahead density forecast of annual inflation.

Empirical Study 1: Forecasting U.S. Inflation

Figure: Qualitative vs. Random Paritioning



(a) LPDR Comparison



(b) Inputs Correlations

This figure shows a comparison between qualitative and random partitioning. Purple line is DRS using qualitative partitioning and black lines are DRS results using partitions chosen at random (computed over random partitions)

Empirical Study 2: Forecasting Industry Stock Returns

Forecast of the annual stock returns for different U.S. industries based on 63 pre-calculated financial ratios and 14 macroeconomic variables.

- 10 industries: Durbl, NonDurbl, Manuf, Energy, HiTech, Health, Shops, Telcm, Utils, Other.

Empirical Study 2: Forecasting Industry Stock Returns

Forecast of the annual stock returns for different U.S. industries based on 63 pre-calculated financial ratios and 14 macroeconomic variables.

- 10 industries: Durbl, NonDurbl, Manuf, Energy, HiTech, Health, Shops, Telcm, Utils, Other.

$J = 10$ groups of predictors

- Value
- Profitability
- Capitalization
- Financial Soundness
- Solvency
- Liquidity
- Efficiency
- Other
- Agg. Financials
- Macro

Empirical Study 2: Forecasting Industry Stock Returns

Forecast of the annual stock returns for different U.S. industries based on 63 pre-calculated financial ratios and 14 macroeconomic variables.

- 10 industries: Durbl, NonDurbl, Manuf, Energy, HiTech, Health, Shops, Telcm, Utils, Other.

Compare against:

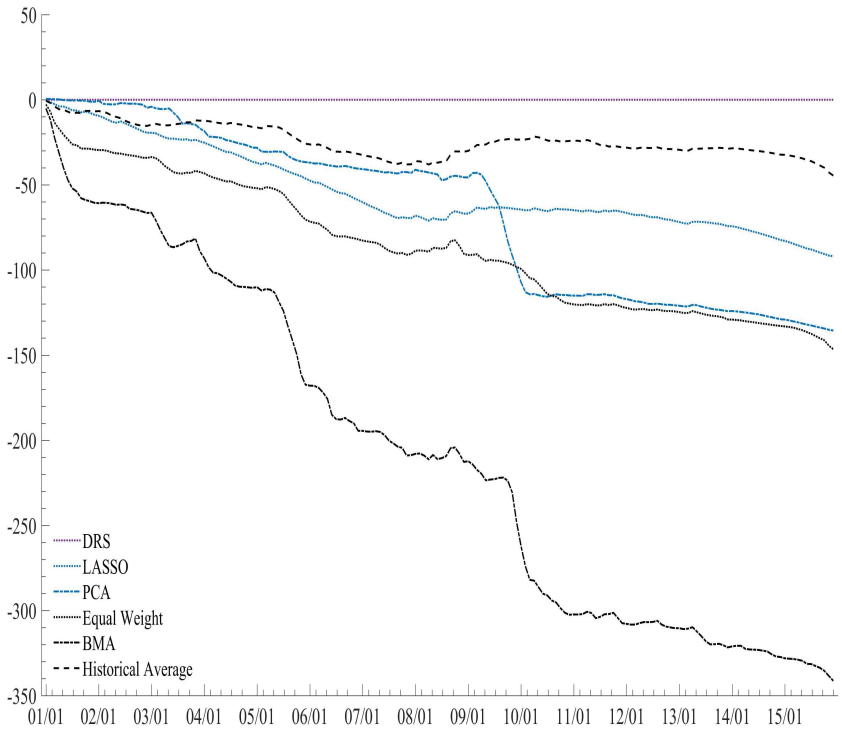
- Sequential BMA
- TVP factor model (k chosen by IC)
- Lasso (leave-one-out cross-validation)
- Linear pooling (equal-weight)
- Sub-group predictions
- Historical Average

$J = 10$ groups of predictors

- Value
- Profitability
- Capitalization
- Financial Soundness
- Solvency
- Liquidity
- Efficiency
- Other
- Agg. Financials
- Macro

Empirical Study 2: Forecasting Industry Stock Returns

Figure: Out-of-sample log predictive density ratio

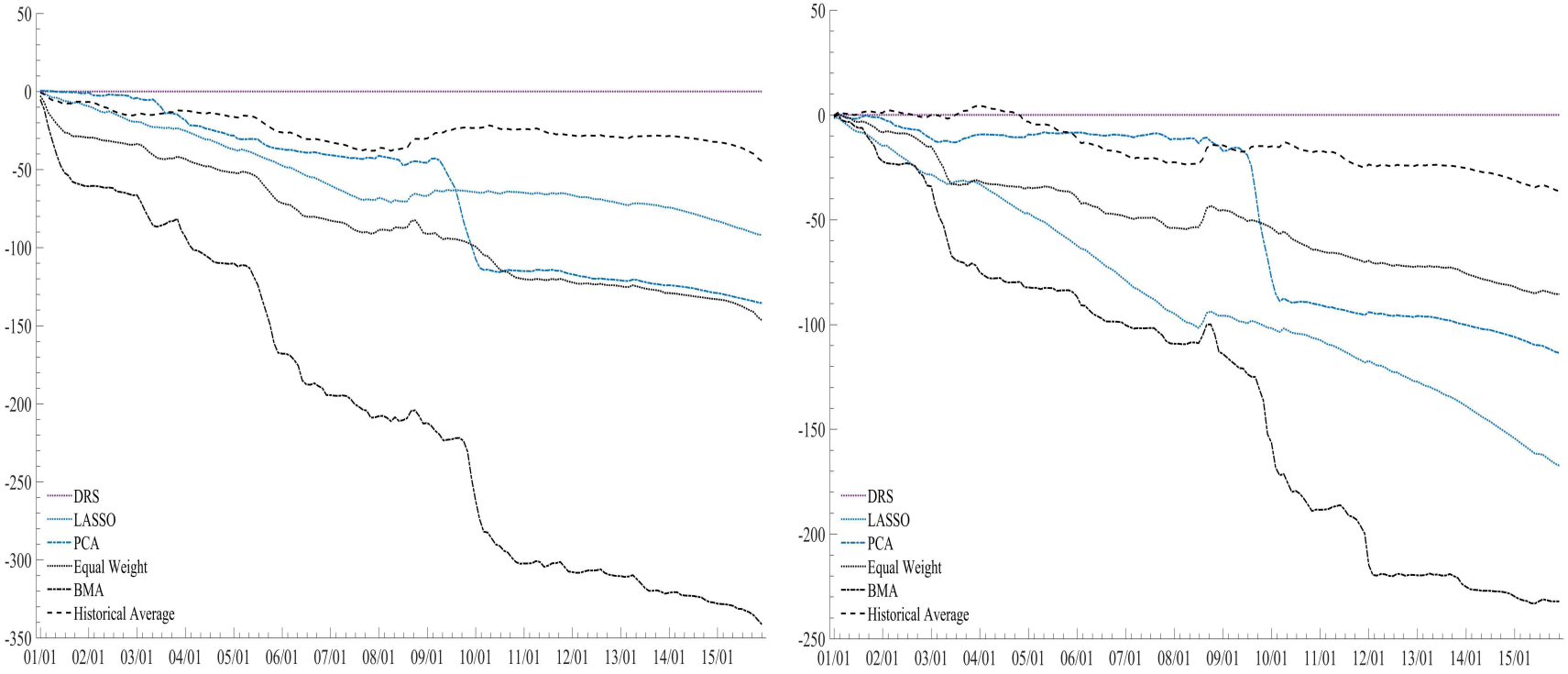


(a) Durable

This figure shows the latent interdependencies across groups of predictive densities used in the recoupling step. Left panel the posterior mean estimates for Durable and right panel for the Manufacturing industry.

Empirical Study 2: Forecasting Industry Stock Returns

Figure: Out-of-sample log predictive density ratio

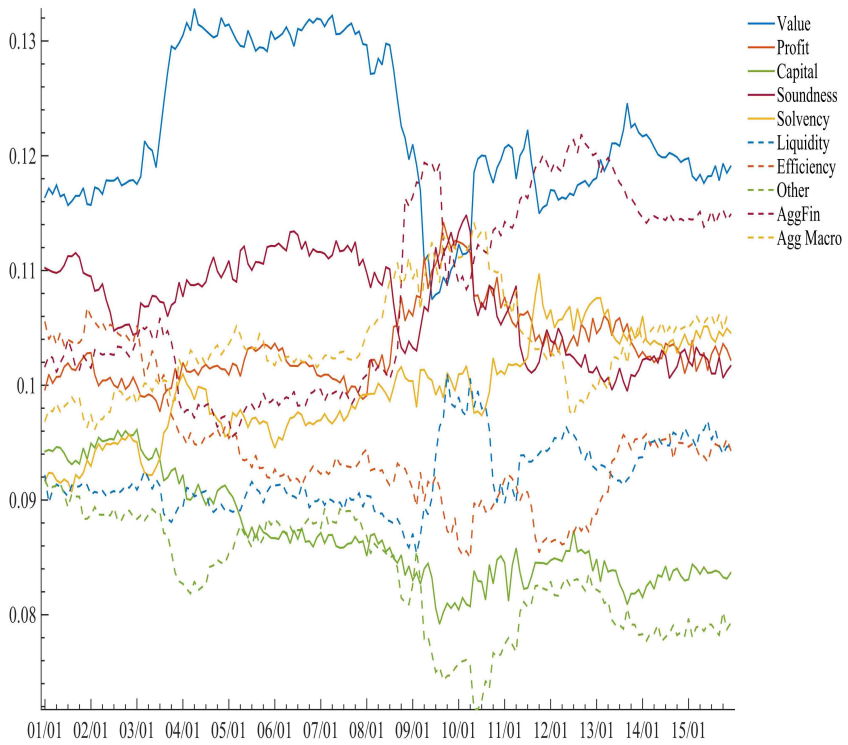


(a) Durable (b) Manufacturing

This figure shows the latent interdependencies across groups of predictive densities used in the recoupling step. Left panel the posterior mean estimates for Durable and right panel for the Manufacturing industry.

Empirical Study 2: Forecasting Industry Stock Returns

Figure: Posterior means of **rescaled** interdependencies.



(a) Durable

This figure shows the latent interdependencies across groups of predictive densities used in the recoupling step. Left panel the posterior mean estimates for Durable and right panel for the Manufacturing industry. The parameters are rescaled between 0 and 1 via logistic transformation.

Empirical Study 2: Forecasting Industry Stock Returns

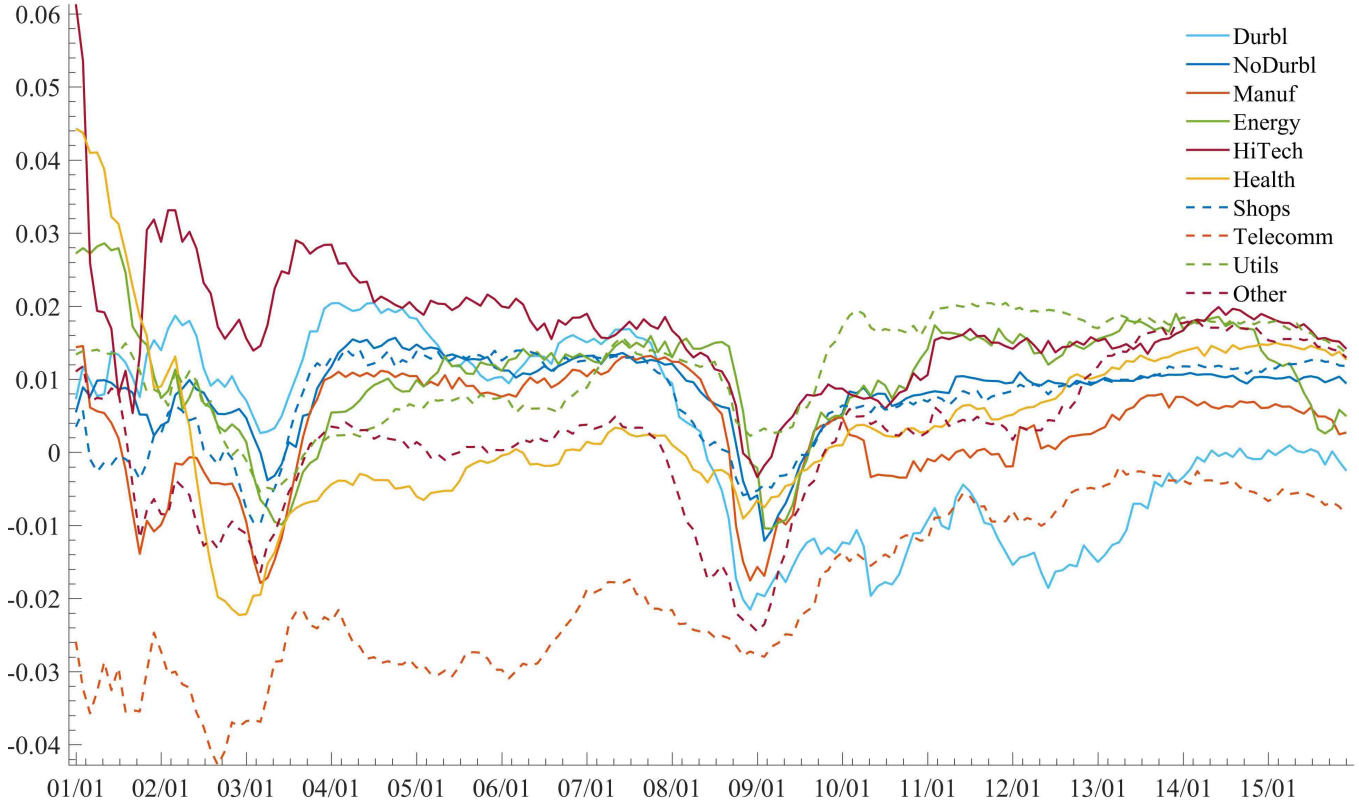
Figure: Posterior means of **rescaled** interdependencies.

(a) Durable (b) Manufacturing

This figure shows the latent interdependencies across groups of predictive densities used in the recoupling step. Left panel the posterior mean estimates for Durable and right panel for the Manufacturing industry. The parameters are rescaled between 0 and 1 via logistic transformation.

Empirical Study 2: Forecasting Industry Stock Returns

Figure: Out-of-Sample Aggregate Bias



This figure shows the dynamics of the out-of-sample predictive bias obtained as the time-varying intercept from the recoupling step of the DRS strategy. The objective function is the one-step ahead density forecast of stock excess returns across different industries. Industry classification is based on 4-digit SIC codes.

Empirical Study 2: Economic Significance

Question: Any economic insight using our predictive strategy?

Empirical Study 2: Economic Significance

Question: Any economic insight using our predictive strategy?

Simplified setting:

- Representative investor with **power utility** and **moderate risk aversion** who wants to allocate wealth between stocks and a risk-less asset.

Empirical Study 2: Economic Significance

Question: Any economic insight using our predictive strategy?

Simplified setting:

- Representative investor with **power utility** and **moderate risk aversion** who wants to allocate wealth between stocks and a risk-less asset.

Two scenarios:

- Short sales are allowed (unconstrained investor).
- Short sales are not allowed, i.e., $w_{it} \geq 0$ (constrained investor).

Empirical Study 2: Economic Significance

Question: Any economic insight using our predictive strategy?

Simplified setting:

- Representative investor with **power utility** and **moderate risk aversion** who wants to allocate wealth between stocks and a risk-less asset.

N.B: Performance is evaluated based on Certainty Equivalent Returns (CERs).

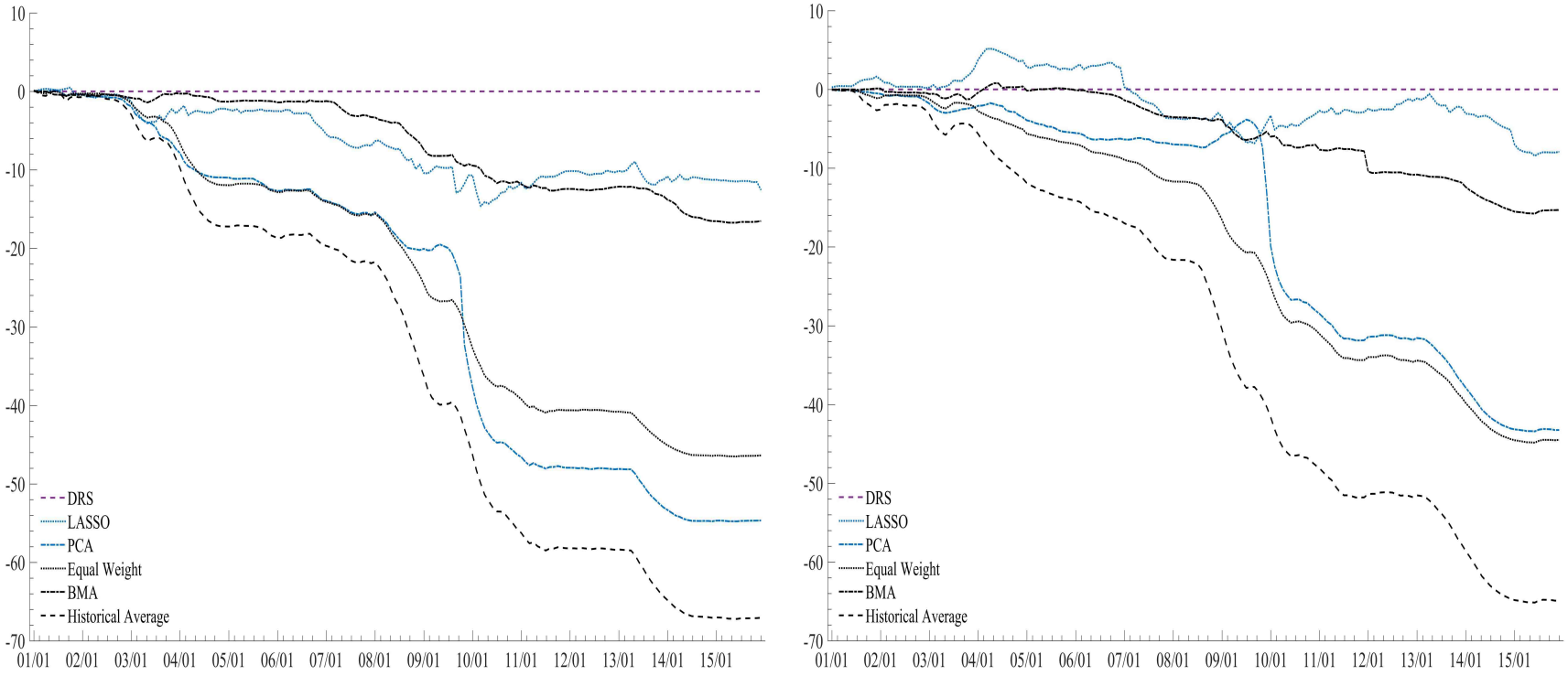
$$CER_{i\tau} = \left[\frac{\hat{U}_{\tau,i}}{\hat{U}_{\tau}} \right]^{\frac{1}{1-\gamma}} - 1,$$

Two scenarios:

- Short sales are allowed (unconstrained investor).
- Short sales are not allowed, i.e., $w_{it} \geq 0$ (constrained investor).

Empirical Study 2: Economic Significance

$$CCER_{it} = \sum_{\tau=1}^{t} \log\left(1 + CER_{i\tau}\right)$$



(a) Durable (b) Manufacturing

Figure: OOS cumulative CER for an **unconstrained** investor

Simulation Study

We consider a simple – yet relevant – simulation study to illustrate the properties of our model.

Data are simulated from

$$y = -2z_1 + 3z_2 + 5z_3 + \epsilon, \quad \epsilon \sim N(0, 0.01),$$
$$z_1 = \frac{1}{3}z_3 + \nu_1, \quad \nu_1 \sim N\left(0, \frac{2}{3}\right),$$
$$z_2 = \frac{1}{5}z_3 + \nu_2, \quad \nu_2 \sim N\left(0, \frac{4}{5}\right),$$
$$z_3 = \nu_3, \quad \nu_3 \sim N(0, 0.01),$$

Characteristics

- Covariates are correlated
- z_3 is omitted
- We assume $\neq$ sub-sampling and procedures.

Results:

- In small samples DRS improves w.r.t other methods by $\approx 60\%$.
- In long samples gains are smaller but persistent.

Conclusions

We propose a **decouple-recouple strategy (DRS) for dynamic predictive regressions** with a (relatively) large number of covariates.

Application to macro and finance:

- DRS **outperforms** typical competitors both **statistically and economically**.

Open issues/next steps:

- Can we group better the predictors?
- Hierarchical structure for larger dimensions?
- How about more sophisticated benchmarks, e.g., DMA?