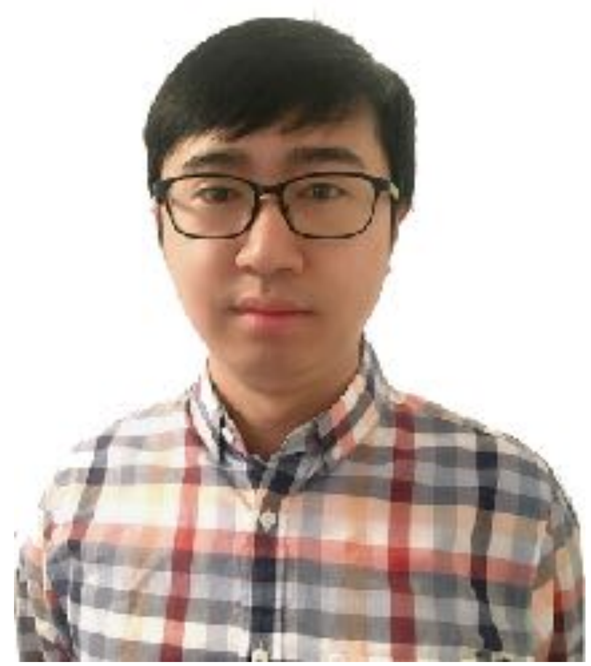
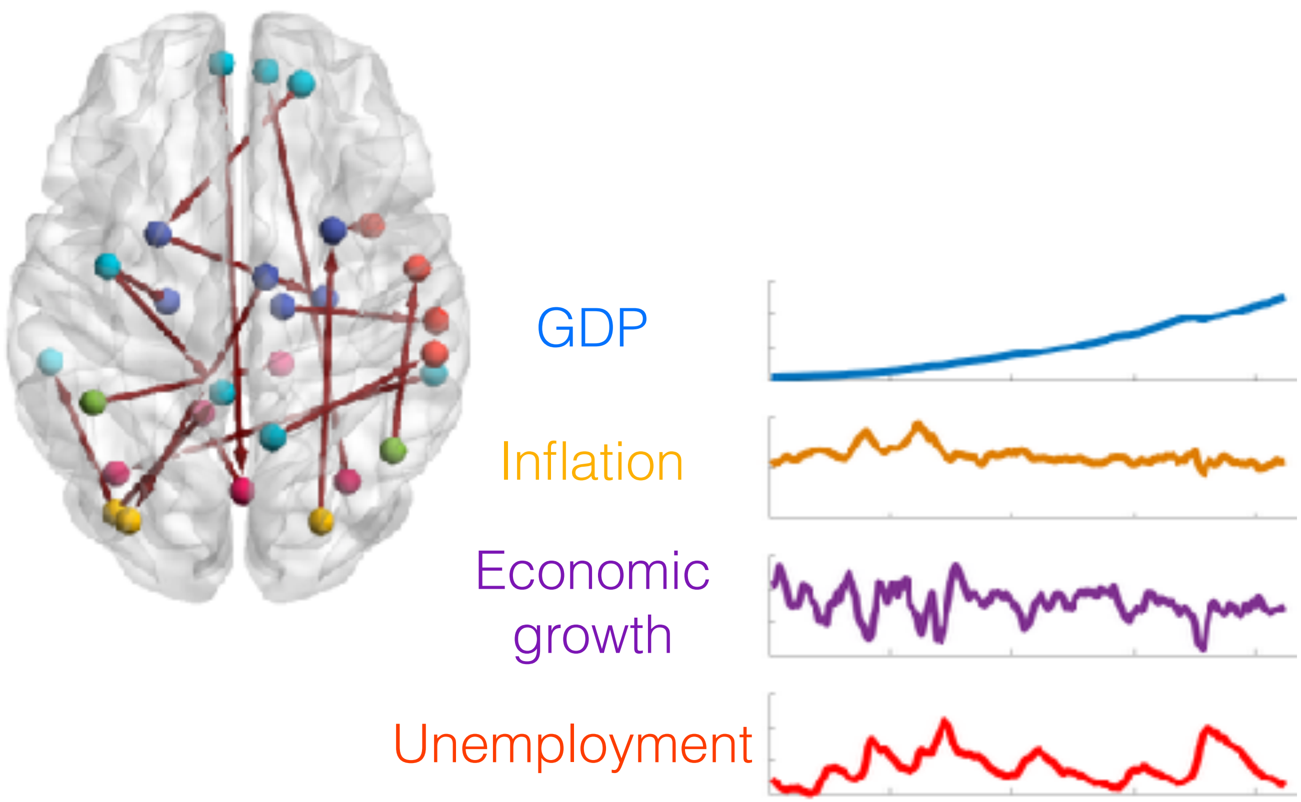
Causal Discovery and Forecasting in Nonstationary Environments with State-Space Models

Biwei Huang[1], Kun Zhang[1], Mingming Gong[1,2], Clark Glymour[1]

1. Department of Philosophy, Carnegie Mellon University
2. Department of Biomedical Informatics, University of Pittsburgh





Two tasks:

1. Identify time-varying causal relations

2. Forecast the values of variables of interest

- Forecasting benefits from causal knowledge
  - Each causal module changes independently
  - Causal knowledge makes the forecasts more interpretable

Time-varying causal model:

$$x_{i,t} = \sum_{x_j \in \mathbb{PA}_i} b_{ij,t} x_{j,t} + e_{i,t}$$

with $e_{i,t} \perp\!\!\!\perp \{b_{ij,t}, x_{j,t}\}_{x_j \in \mathbb{PA}_i}$.

- $b_{ij,t}$ and $\sigma^2_{i,t}$ change over time

$$b_{ij,t} = \alpha_{ij,0} + \sum_{p=1}^{p_l} \alpha_{ij,p} b_{ij,t-p} + \epsilon_{ij,t},$$

$$h_{i,t} = \beta_{i,0} + \sum_{q=1}^{q_l} \beta_{i,q} h_{i,t-q} + \eta_{i,t},$$

where $h_{i,t} = \log \sigma^2_{i,t}$.

$e_{i,t}, \epsilon_{ij,t}, \eta_{i,t}$

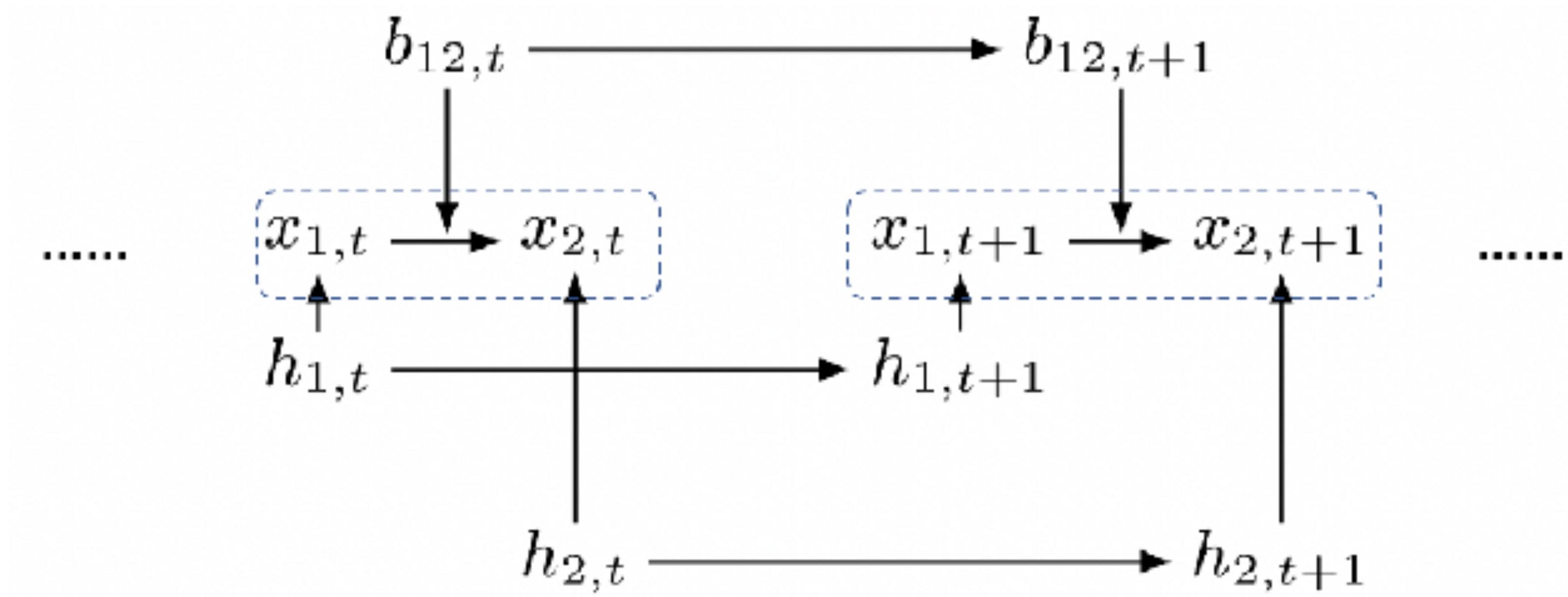
- Gaussian
- Mixture of Gaussian

Causal Model, Identifiability, and Estimation

- **Goal**: Find time-varying causal relations & make prediction
- Causal relations change over time
  - **Model**: causal coefficients modeled by *autoregressive models*
- **Identifiability**: The causal model identifiable if the underlying causal structure is acyclic
- **Model Estimation**: A specific nonlinear state-space model
  - Estimated by Stochastic approximation EM with Conditional particle filter

Forecasting with time-varying causal model

- Treat forecasting as a Bayesian inference problem in the causal model

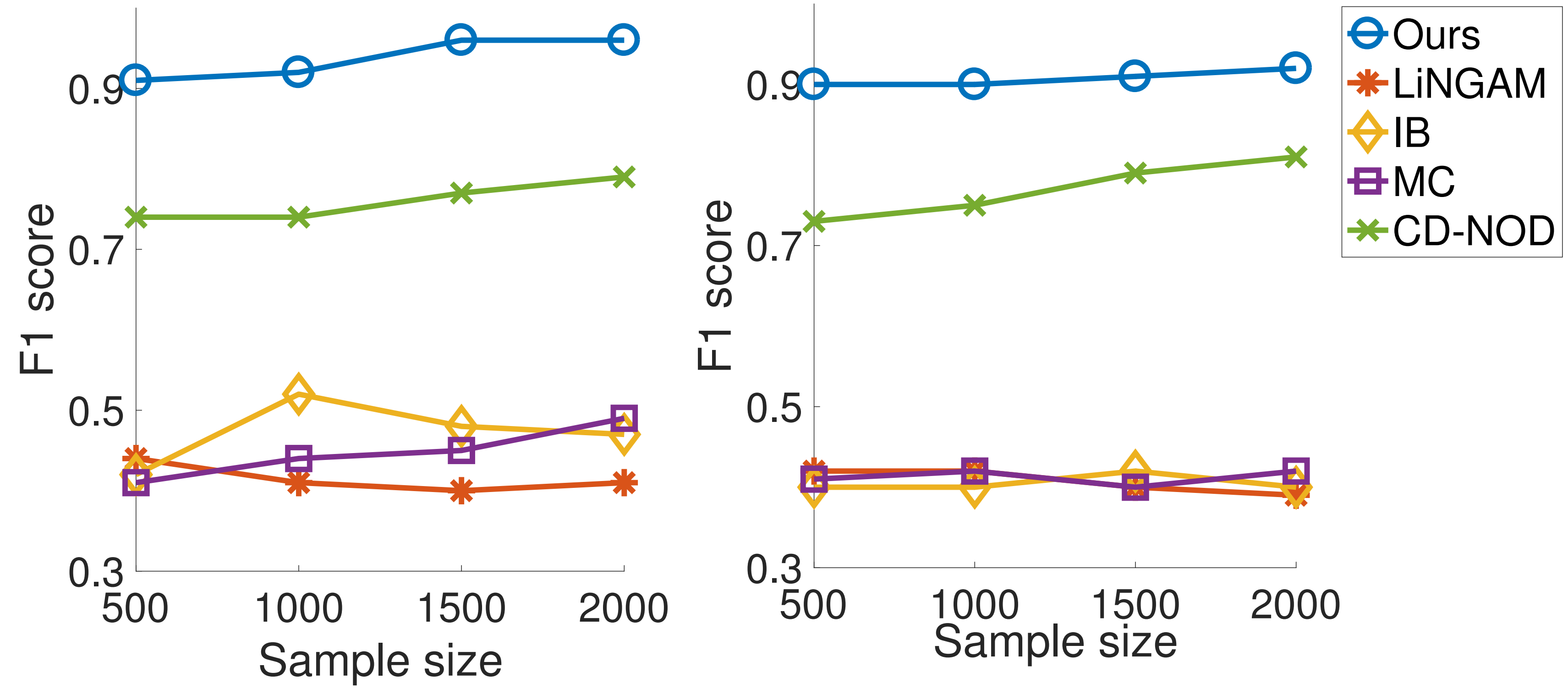


- Metropolis-Hastings to forecast Y_{T+1}

$$p(Y_{T+1}|\mathcal{M}_{Y,T+1}) \propto p(Y_{T+1}|\mathcal{P}_{Y,T+1}) \prod_{\tilde{X}_{C_i} \in \mathcal{C}_Y} p(\tilde{X}_{C_i,T+1}|\mathcal{P}_{C_i,T+1})$$

Causal discovery:

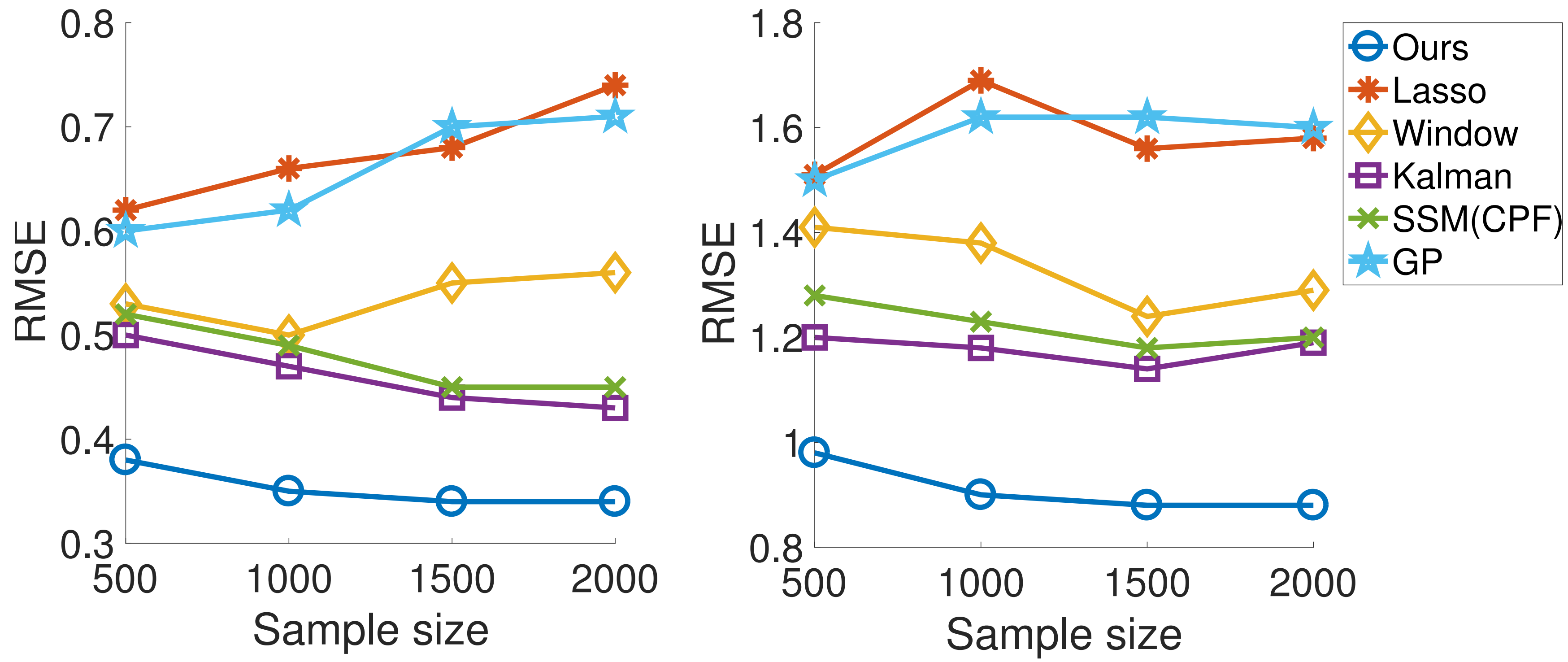
Ours: highest F1 score!



(a) Only b changes (b) Both b and σ^2 change

Forecasting:

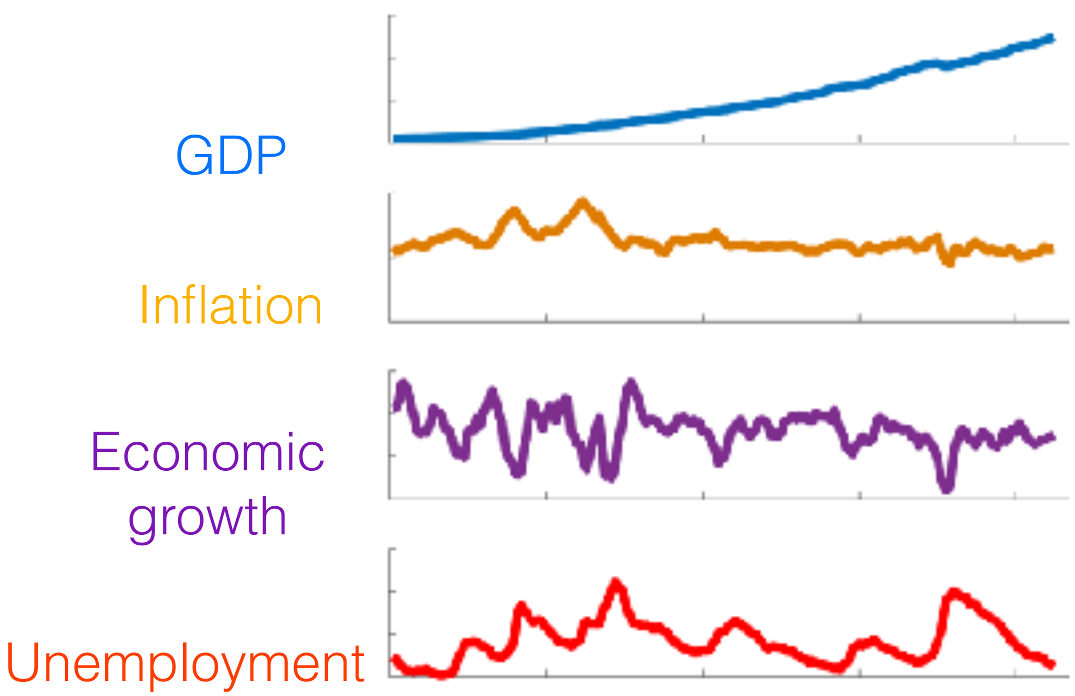
Ours: lowest RMSE!

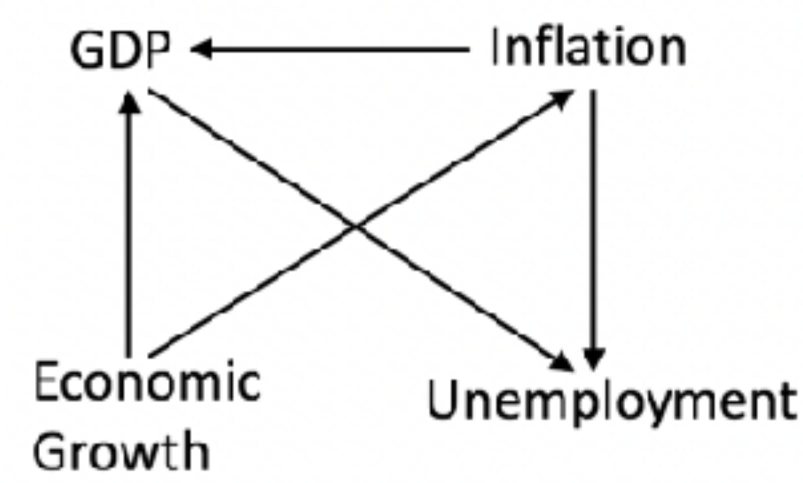


(a) Only b changes (b) Both b and σ^2 change

Macroeconomics data

(quarterly data, 1965-2017, USA)





Methods	RMSE	Methods	RMSE
Ours	**0.32**	Lasso	0.38
Kalman filtering	0.42	Window Lasso	0.37
SSM (CPF)	0.43	GP	0.37

RMSE of the forecasts on inflation (2007 - 2017).

Poster #73

Conclusion

- A unified framework for causal discovery and forecasting
- Establish the identifiability results, even when data is conditional Gaussian

Future work

- Improve the scalability
- Nonlinear causal relationships, partially observable processes, and causal models with instantaneous cycles…