Econometrics and Data Science: Research Projects

Siem Jan Koopman, prof. dr. in Econometrics

Topic: **Presentation EDS Group**

School of Business and Economics, VU Amsterdam
Ksandr Live XL– June 17, 2019

VU VRIJE UNIVERSITEIT AMSTERDAM

Outline

Econometrics and Data Science: The EDS Group

EDS Group

- **EDS Group** operates within the

  School of Business and Economics, VU Amsterdam
  Department of Econometrics and Operations Research

- **EDS Group** consists of 20+ fte
  - 12 fte research staff (senior, junior, tenure-track)
  - 4 part-time researchers (business partners, 1 fte)
  - 2 fte post-docs (external funding)
  - 6 PhD students (core 4 fte)
  - 2 fte teaching staff

Econometrics and Data Science

VU VRIJE UNIVERSITEIT AMSTERDAM

- **Data Science**:

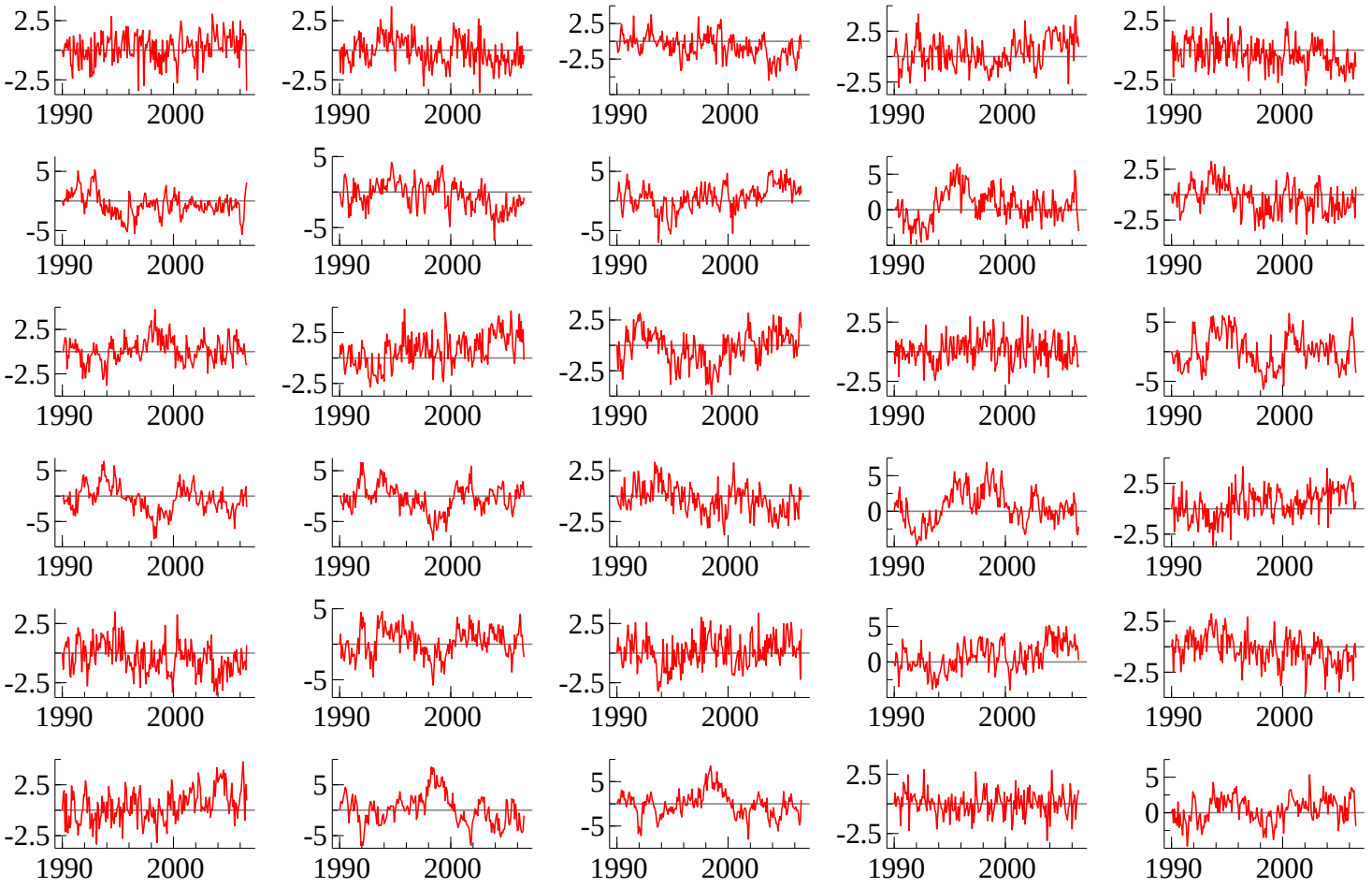
  - Multivariate Statistical Methods
  - Network structures
  - Classification
  - Regression
  - Unsupervised Statistical Learning (k-means, principal components)
  - Random Forests

- **Econometrics**

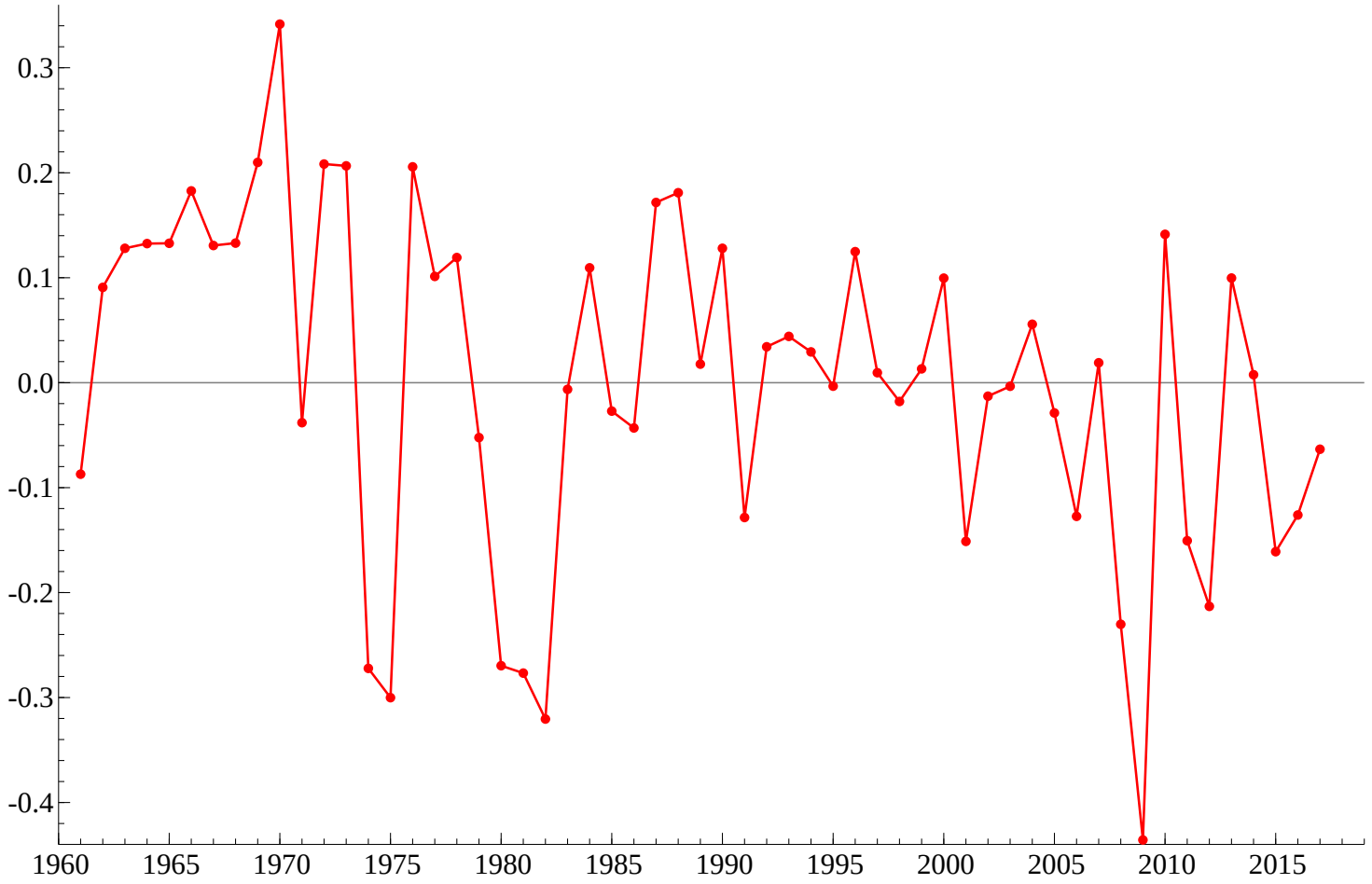
  - Time Series and Dynamic Econometrics
  - Prediction and Forecasting
  - Causal Structural Modelling
  - Mixed-Frequency, Messy and Noisy Data

Factors and Prediction

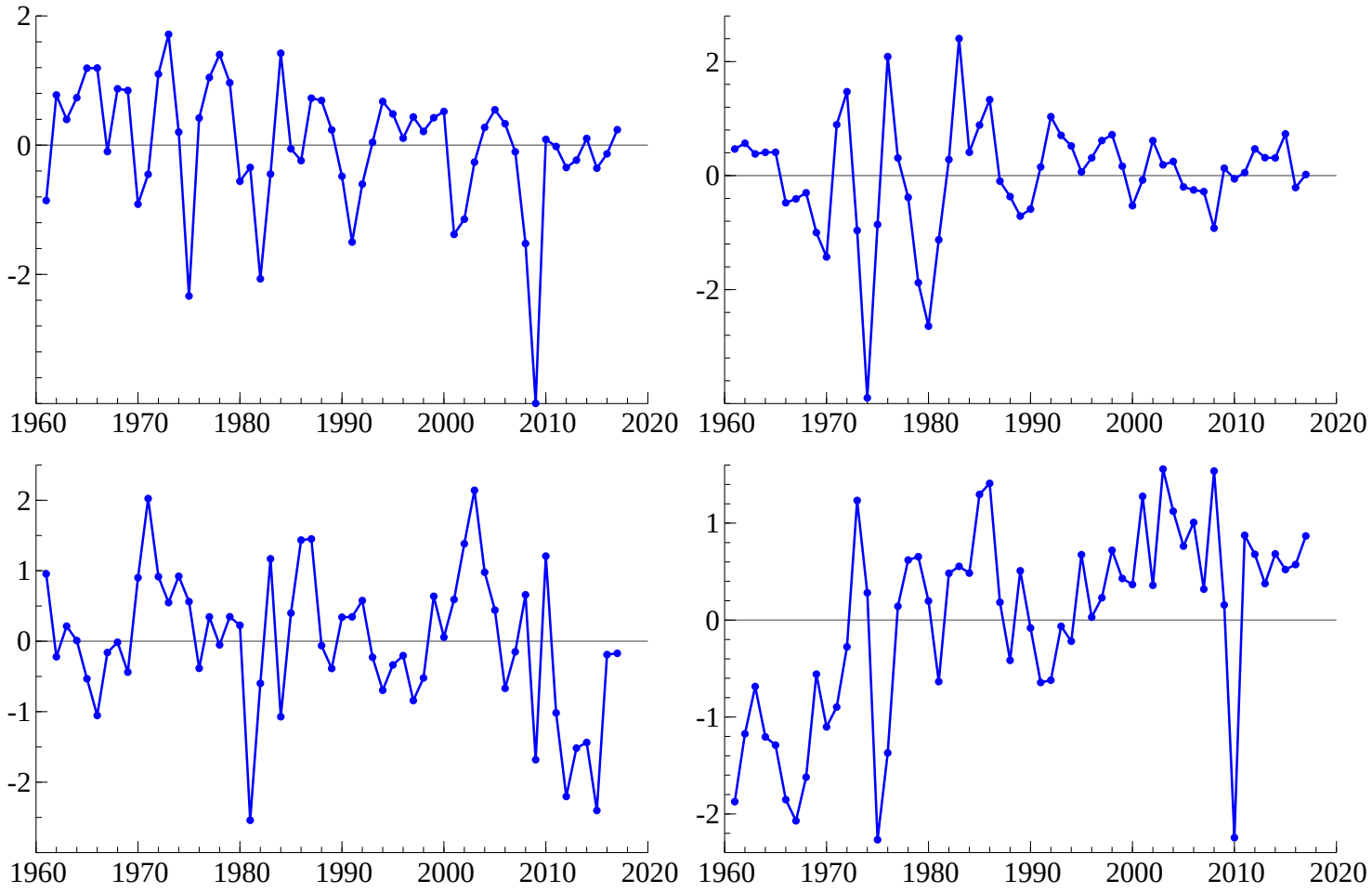
Big Data



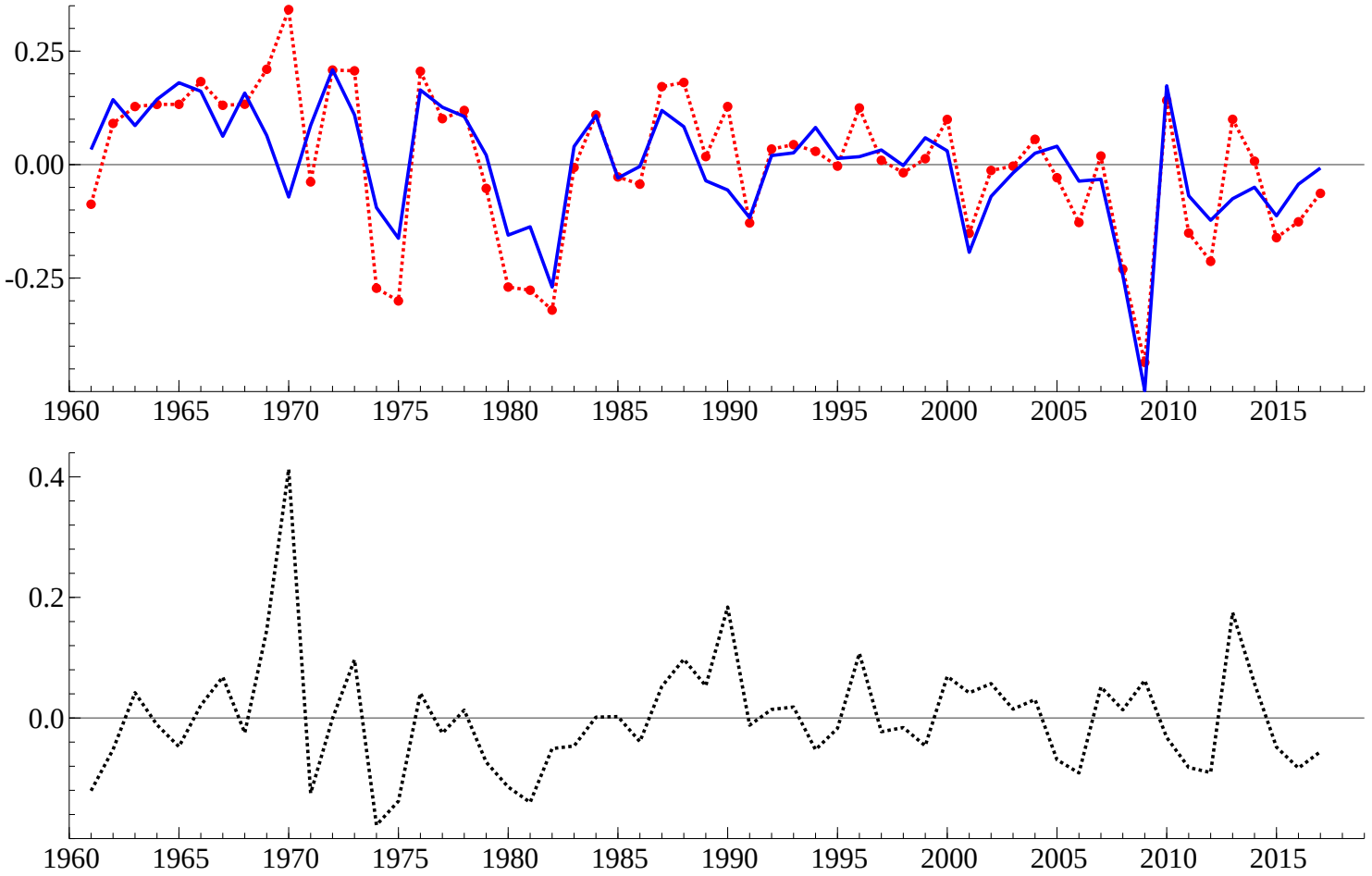
One series of interest



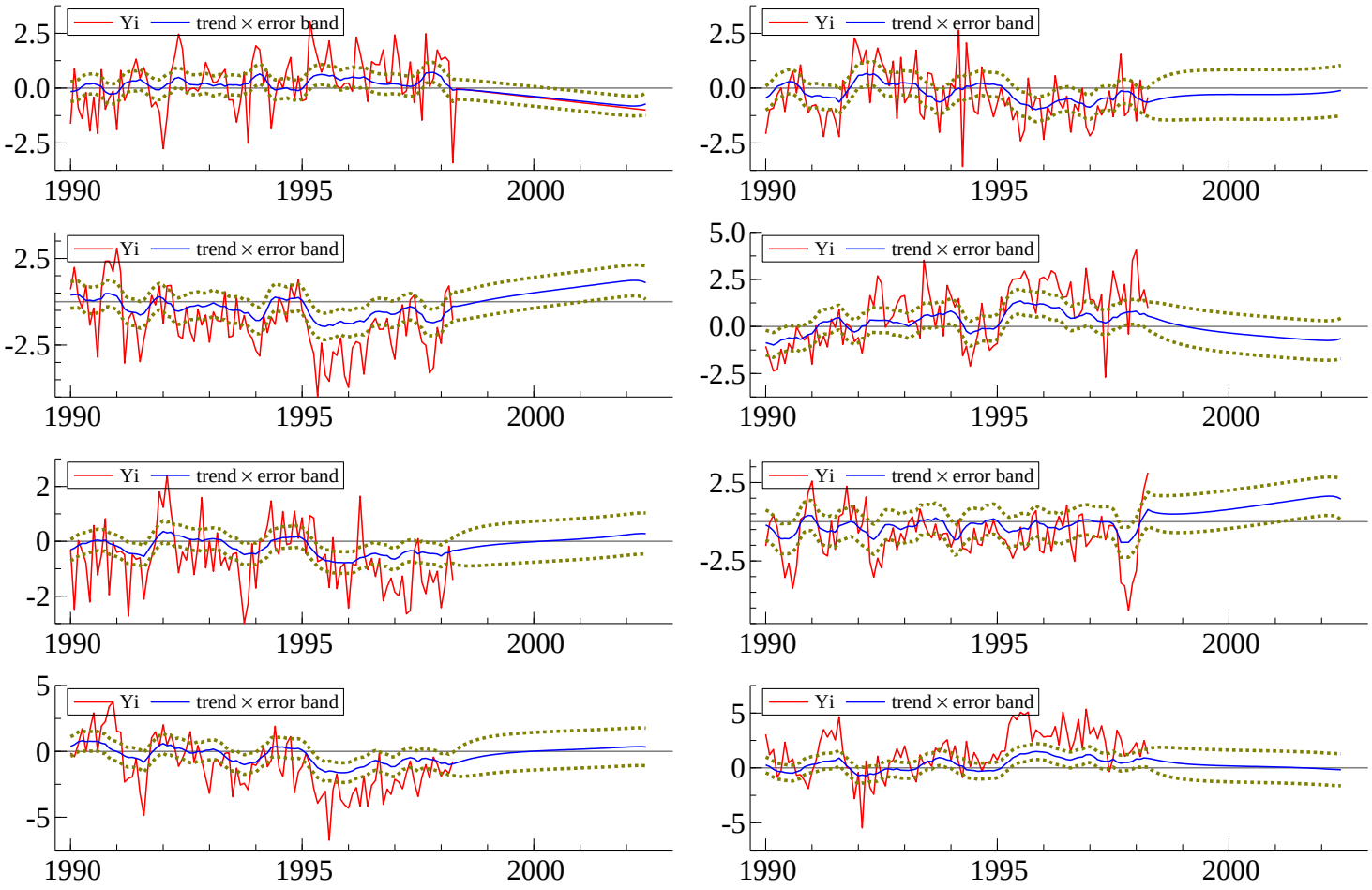
Factors


2
0
-2
1960
1970
1980
1990
2000
2010
2020
2
0
-2
1960
1970
1980
1990
2000
2010
2020
2
1
0
-1
-2
1960
1970
1980
1990
2000
2010
2020
1
0
-1
-2
1960
1970
1980
1990
2000
2010
2020

Fit


0.25
0.00
-0.25
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
0.4
0.2
0.0
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015

Prediction


Yi
trend × error band
2.5
0.0
-2.5
1990
1995
2000
Yi
trend × error band
2.5
0.0
-2.5
1990
1995
2000
Yi
trend × error band
2.5
-2.5
1990
1995
2000
Yi
trend × error band
5.0
2.5
0.0
-2.5
1990
1995
2000
Yi
trend × error band
2
0
-2
1990
1995
2000
Yi
trend × error band
2.5
-2.5
1990
1995
2000
Yi
trend × error band
5
0
-5
1990
1995
2000
Yi
trend × error band
5
0
-5
1990
1995
2000

Mixed, Messy and Noisy Data

Signal and Noise

Key challenge: from noisy indicators towards SIGNAL EXTRACTION

We are the experts in effective **SIGNAL EXTRACTION**:

- Building predictive models from data with
- different frequencies, different features, missing entries, outliers, etc.
- Using Score-Driven framework:

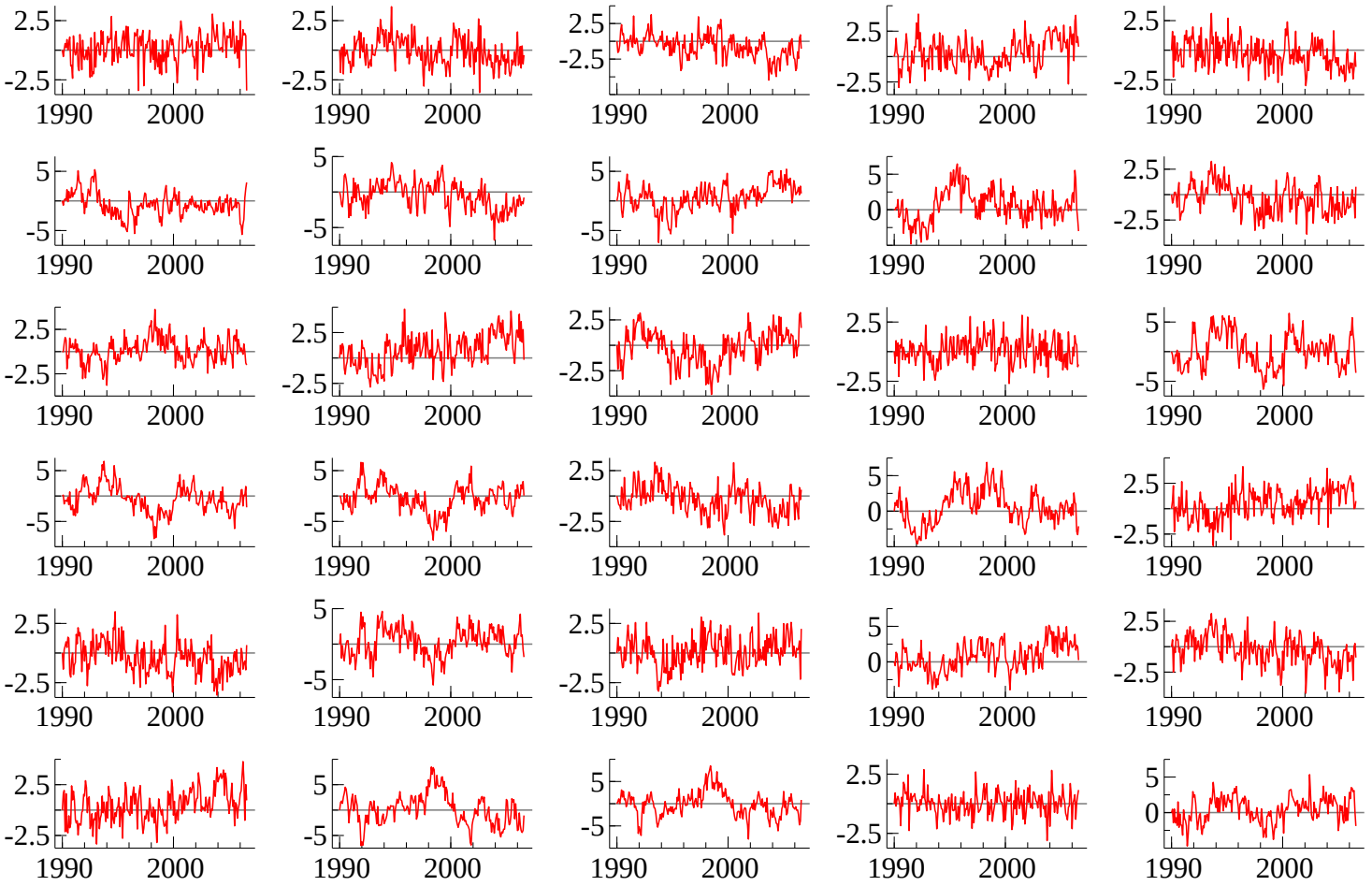
$$y_{it} \sim p_i(y_{it}|\mathcal{F}_{t-1}, \theta_t; \psi), \qquad \theta_t = Z_t \alpha_t$$

- with data y_{it}, distribution $p(y_{it}; \cdot)$, past data $\mathcal{F}_{t-1}$, signal θ_t, parameter vector ψ and state vector α_t with dynamic updating

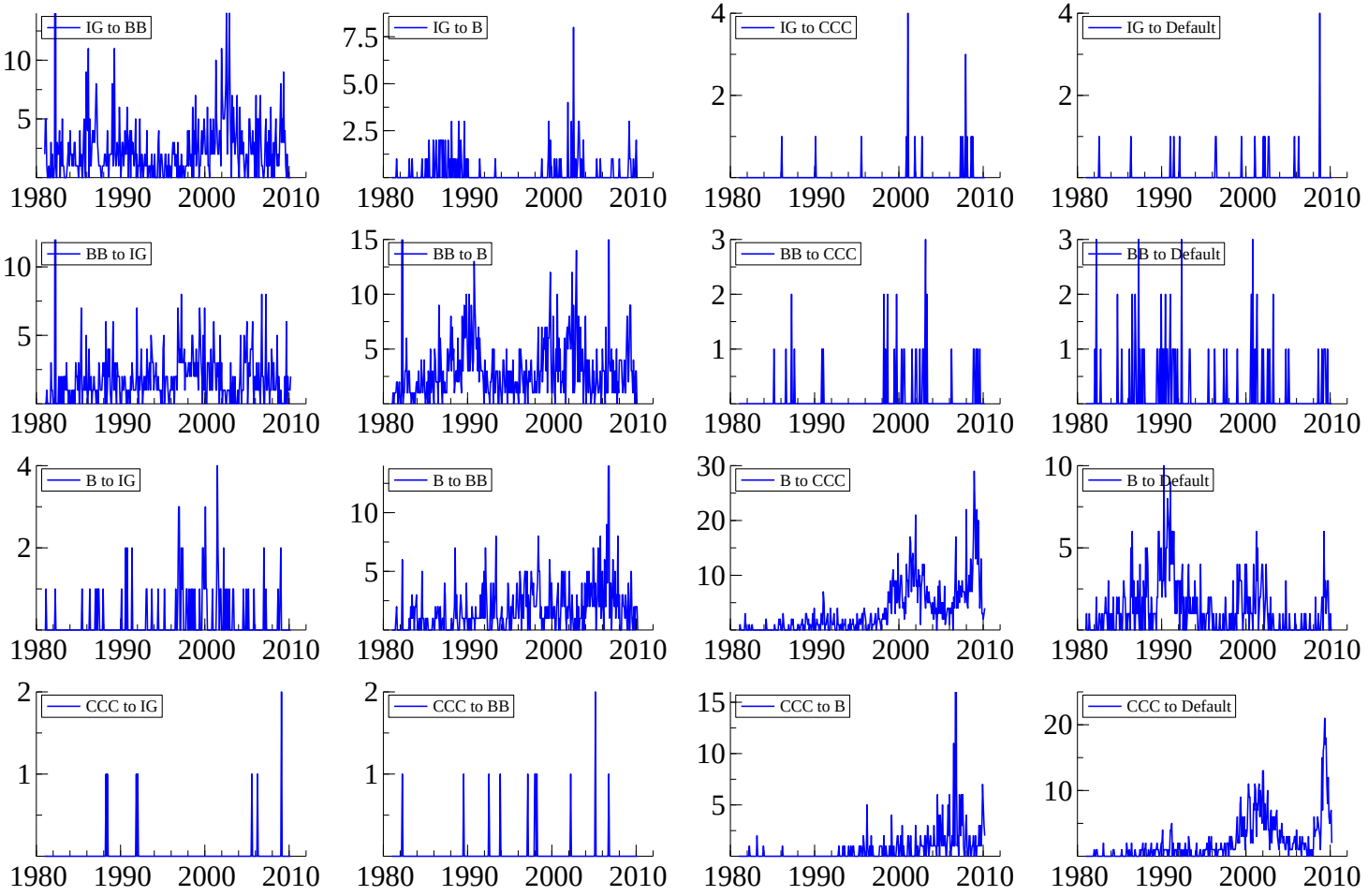
$$\alpha_{t+1} = \omega + \beta \alpha_t + \delta \nabla_t$$

where ∇_t is the innovation (score function) and ω, β, δ are coefficients.

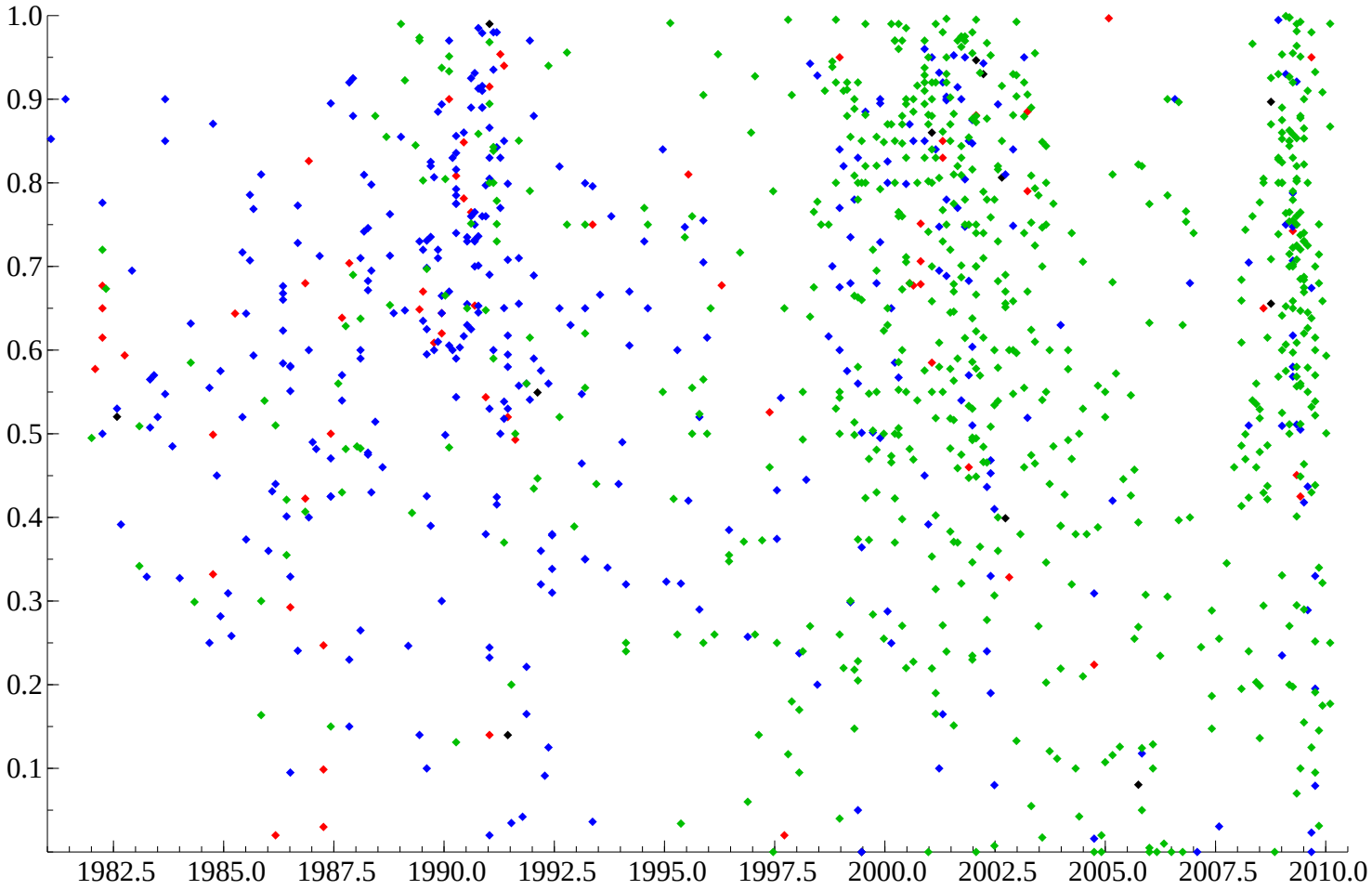
Risk Measure 1: Macro Data



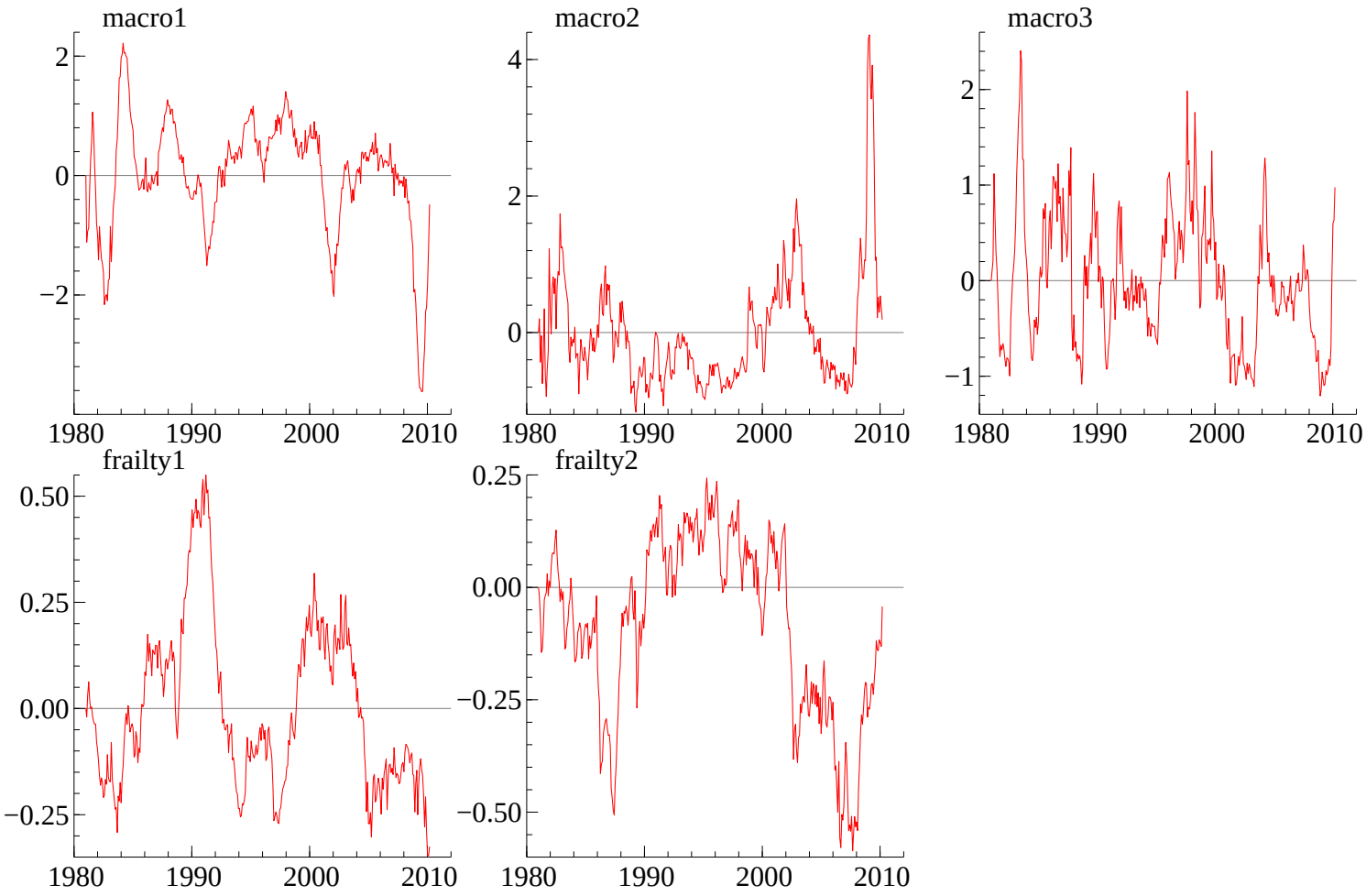
Risk Measure 2: Failure Counts



Risk Measure 3: Risk Probabilities



Risk Signals

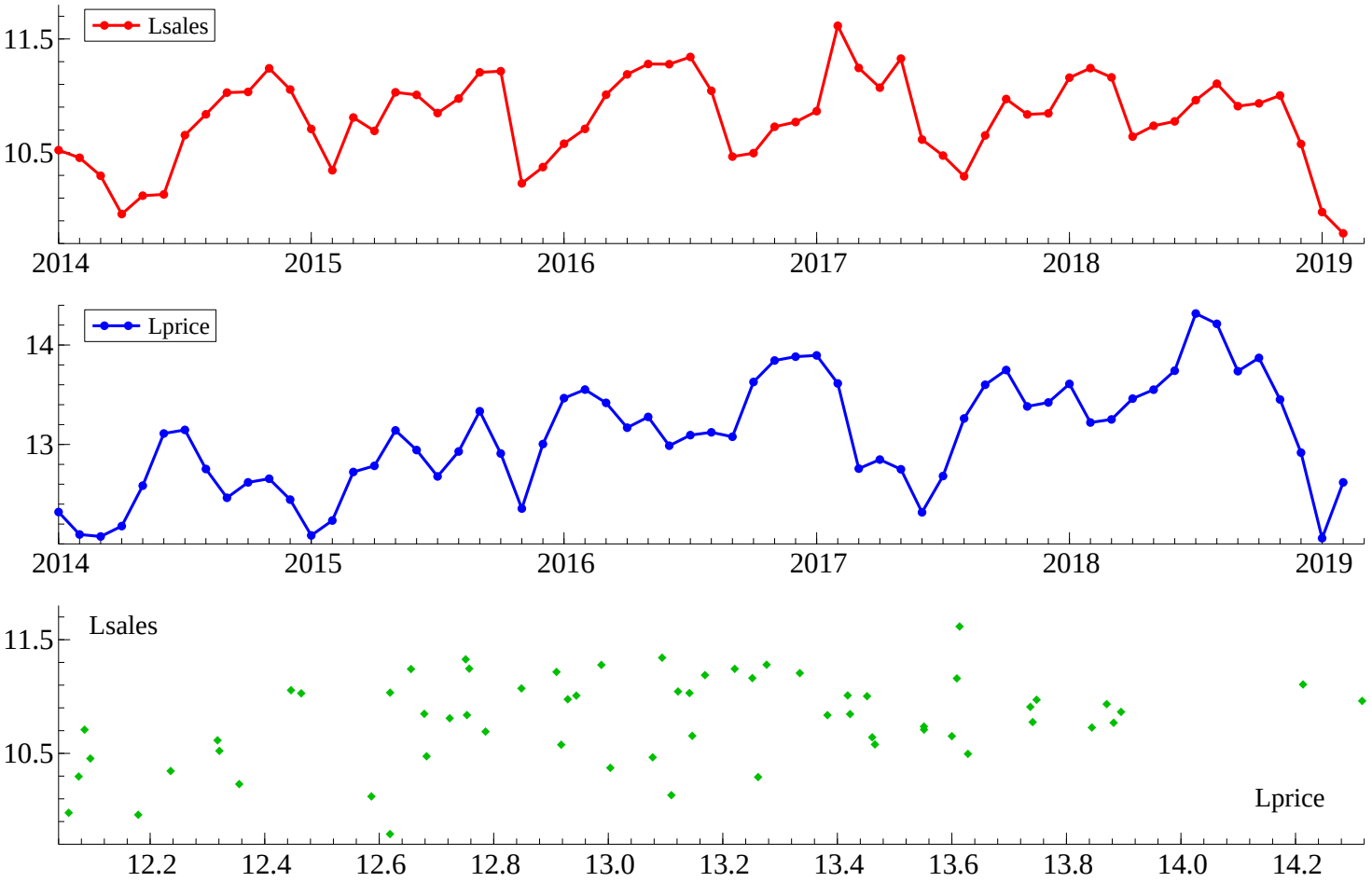


Econometric Causal Models

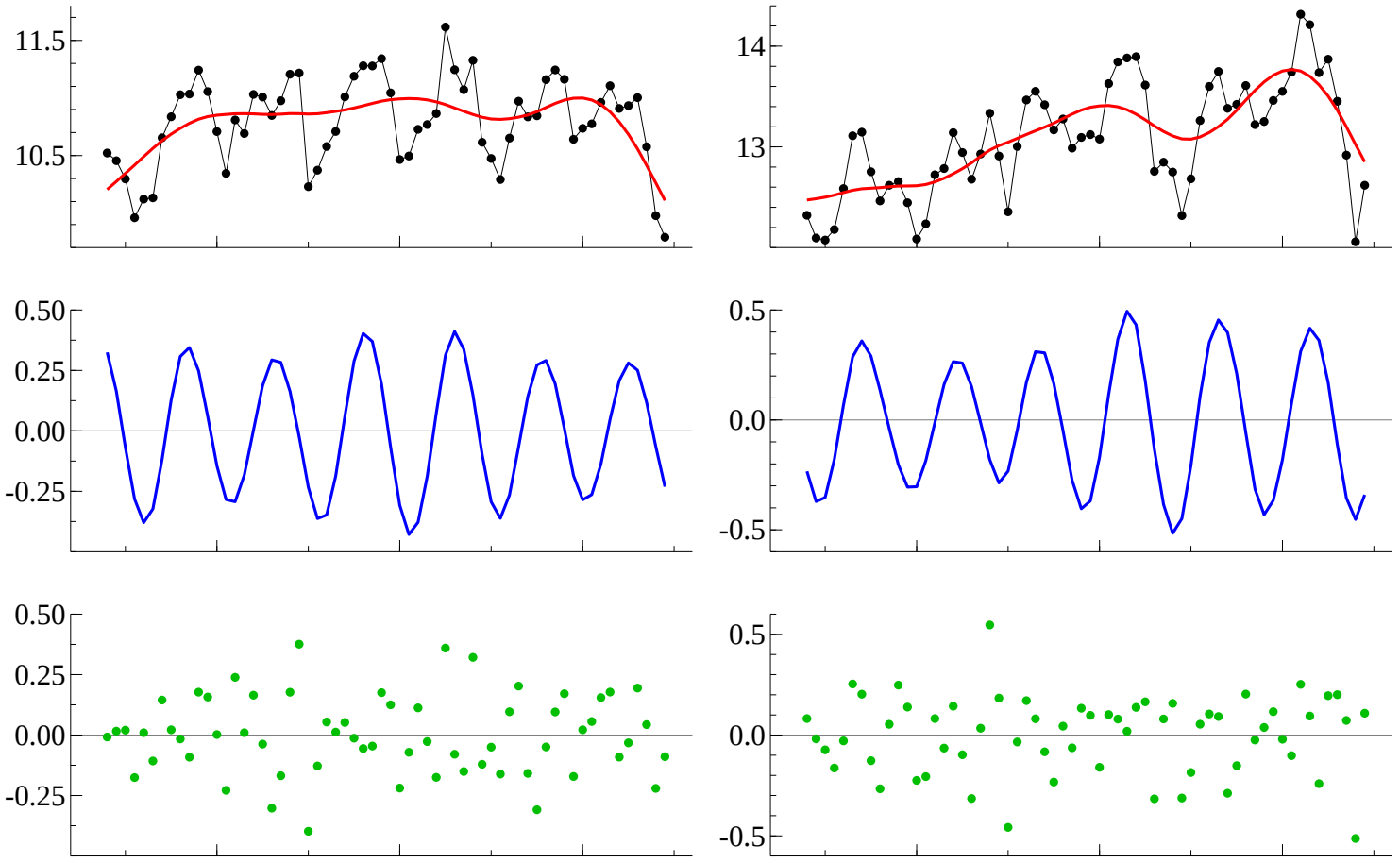
Econometrics and Causality

- Data Science Methods are highly effective in Data Exploration, Visualisation, Correlations
- Correlation versus Causality: structural econometric models
- Classic example is sales/consumption versus price/inflation

Sales and Price: positive correlation ?



Sales and Price: structural analysis



- In a simultaneous analysis: common long-term and short-term features are recognised
- The temporary (medium-term) effects are extracted
- The medium-term effects are negatively related by -0.5: price elasticity

Conclusion

- EDS Group from VU Amsterdam
- Broad scope on Econometrics and Data Science Methods
- Structural view on Data Analysis, Modelling and Prediction