Identifying Monetary Policy Shocks: A Natural Language Approach

S. Borağan Aruoba **Thomas Drechsel**

July 2022

MOTIVATION

- Goal: study the effects of monetary policy i_t on macro variables $\boldsymbol{Y_t}$
- Problem: i_t endogenously reacts to $\boldsymbol{Y_t}$
- Formalize problem:

$$i_t = f(\Omega_t) + \varepsilon_t$$

  - Ω_t: information set of central bank, contains $\boldsymbol{Y_t}$
  - $f(\cdot)$: systematic conduct of monetary policy
  - ε_t: monetary policy shock

MOTIVATION

- Idea of Romer and Romer (2004)
  - Run linear regression

$$\Delta i_t = \alpha + \beta i_{t-1} + \boldsymbol{\gamma} \boldsymbol{X_t} + \varepsilon_t^{RR}$$

    $\boldsymbol{X_t}$ contains forecasts from documents prepared for FOMC ("Greenbooks")
  - With residuals $\hat{\varepsilon}_t^{RR}$, construct IRFs of $\boldsymbol{Y_t}$

MOTIVATION

- Idea of Romer and Romer (2004)
  - Run linear regression
    $$\Delta i_t = \alpha + \beta i_{t-1} + \boldsymbol{\gamma} \boldsymbol{X_t} + \varepsilon_t^{RR}$$
    $\boldsymbol{X_t}$ contains forecasts from documents prepared for FOMC ("Greenbooks")
  - With residuals $\hat{\varepsilon}_t^{RR}$, construct IRFs of $\boldsymbol{Y_t}$
- Key assumptions:
  1. Forecasts of Fed economists are good approximation of information set Ω_t
  2. Linear specification is good approximation of systematic policy $f(\cdot)$

MAIN IDEA OF THIS PAPER

- Revives the idea of Romer and Romer (2004)
- Proposes new approach using **natural language processing** & **machine learning**
  - Process *more* information in Fed documents, including *human language*
  - Allow for nonlinearities in how information translates into policy
- Shows that additional information + nonlinearities more plausibly represent $f(\Omega_t)$
- Studies the effects of monetary policy shocks on macro variables

PREVIEW OF OUR METHODOLOGY

- **Natural language processing:**
  - Process language in thousands of pages of text prepared for FOMC meetings
  - Obtain sentiment indicators for economic concepts that are discussed
- **Machine learning:**
  - Include numerical forecasts and sentiment indicators in a regression
  - For both, include linear and nonlinear terms $\rightarrow$ hundreds of potential regressors
  - Apply ridge regression as *dense* ML technique

PREVIEW OF FINDINGS 1/2

- Contribution of systematic vs. exogenous changes in monetary policy
  - Original Romer-Romer regression: $R^2 = 0.5$; implies 50% of ΔFFR are shocks
  - Our approach $R^2 = 0.75$; implies half of original RR shocks are endogenous
  - Wider side of forecasts, inclusion of human language, and nonlinearities all contribute
  - Additional information not useful: committee composition, information in transcripts
- Inspecting the drivers of systematic changes in monetary policy
  - Mostly real activity sentiments and forecasts
  - Limited role for sentiments around price and financial variables

PREVIEW OF FINDINGS 2/2

- What are monetary policy shocks?
  - FOMC makes decisions not directly related to economic outlook
  - E.g. based on long-run credibility concerns
- Effects of monetary policy shocks
  - Estimated shocks give theoretically consistent IRFs of standard variables
  - Not the case for shocks estimated with original Romer-Romer specification

$$i\uparrow \Rightarrow \; Y\downarrow \;\; P\downarrow \;\; EBP\uparrow \;\; SP500\downarrow$$

METHODOLOGY

1. PROCESS RAW TEXT

- Download documents associated with scheduled FOMC meetings
  - Main focus: Beigebook & Tealbook A (for earlier dates: Red- & Greenbooks)
- Start with the meeting on October 5, 1982, which is when the Fed started targeting the Fed Funds Rate as their policy tool, see Thornton (2006)
- End with latest available meeting in December 2016 $\Rightarrow$ 276 FOMC meetings
- Some of subsequent analysis runs until 2008:10: last unconstrained policy change before ZLB, includes 210 FOMC meetings

2. IDENTIFY ECONOMIC CONCEPTS

- After cleaning steps, we store all singles, doubles, and triples
  - "... consumer price inflation ..." gives a triple, two doubles and three singles
  - "... inflation and economic activity ..." gives us three singles and one double
  - "... for inflation. Activity on the other hand..." gives us three singles
- Select most frequently discussed economic concepts
  - This step involves human judgment
- Combine/exclude overlapping concepts Details
- Final list amounts to 296 economic concepts

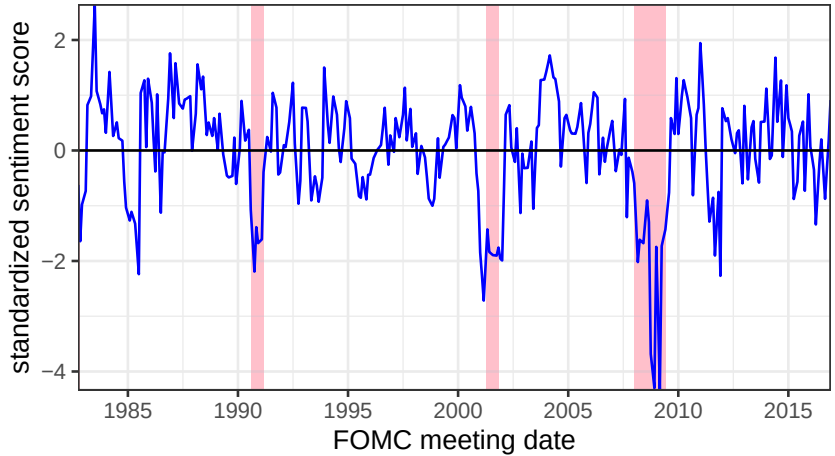
3. MOST FREQUENT ECONOMIC CONCEPTS

3. CONSTRUCT SENTIMENT

- Apply a method inspired by Hassan, Hollander, van Lent, and Tahoun (2020)
- Consider the 10 words mentioned before and after each concept's appearance
- Check whether these are words associated with positive or negative sentiment
  - Use classification based on enhanced version of Loughran and McDonald (2011)
- Each positive word gives a score of +1 and each negative word of -1
- Sum up the scores within a meeting, and scale by the total number of words
- Using sentences instead of +/- 10-word windows gives very similar indicators

Dictionary example

EXAMPLE: SENTIMENT AROUND "ECONOMIC ACTIVITY"



More

4. RUN RIDGE REGRESSION

$$\Delta i_t = \alpha + \beta i_{t-1} + \Gamma(\widetilde{\boldsymbol{X}}_t, \boldsymbol{Z}_t) + \varepsilon_t^*$$

- $\widetilde{\boldsymbol{X}}_t$: numerical forecasts: with all variables, lags, differencing $\rightarrow$ 132 time series
- $\boldsymbol{Z_t}$: sentiment indicators $\rightarrow$ 296 time series
- $\Gamma(\cdot)$ captures non-linearity $\rightarrow$ implement as linear-quadratic specification
- Problem is "curse of dimentionality"
  - In above setting, 858 variables on the right hand side
  - Before ZLB, 210 observations
- Solution: ridge regression

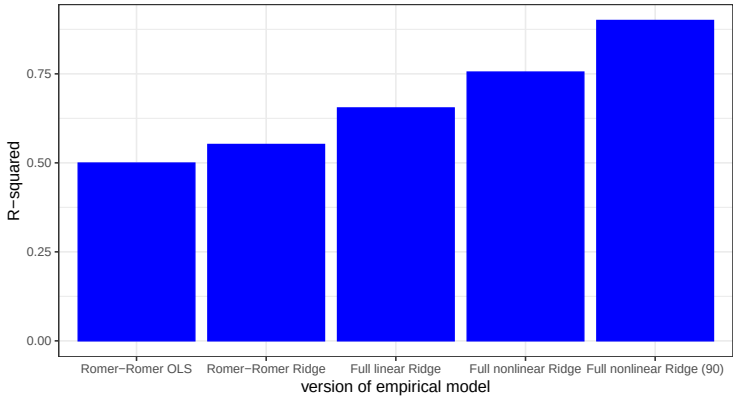
4. RUN RIDGE REGRESSION

- While OLS minimizes RSS, Ridge minimizes $RSS + \lambda \sum_n^N \beta_n^2$
  - Bayesian interpretation: set a normal prior with scale λ on $\boldsymbol{\beta}$
  - Alternatives, e.g. LASSO, which we try in robustness checks

- How to set ridge tuning parameter λ?
  - **Option 1**: optimally choose based on k-fold cross-validation CV
  - **Option 2**: a priori restriction on contribution of systematic policy
    - "*Even the harshest critics of monetary authorities would not maintain policy decisions are unrelated to the economy.*" (Leeper, Sims, and Zha, 1996)
    - We suggest restriction that imposes $R^2 = 0.9$

RESULTS OF THE IDENTIFICATION PROCEDURE

R^2 ACROSS DIFFERENT REGRESSION MODELS



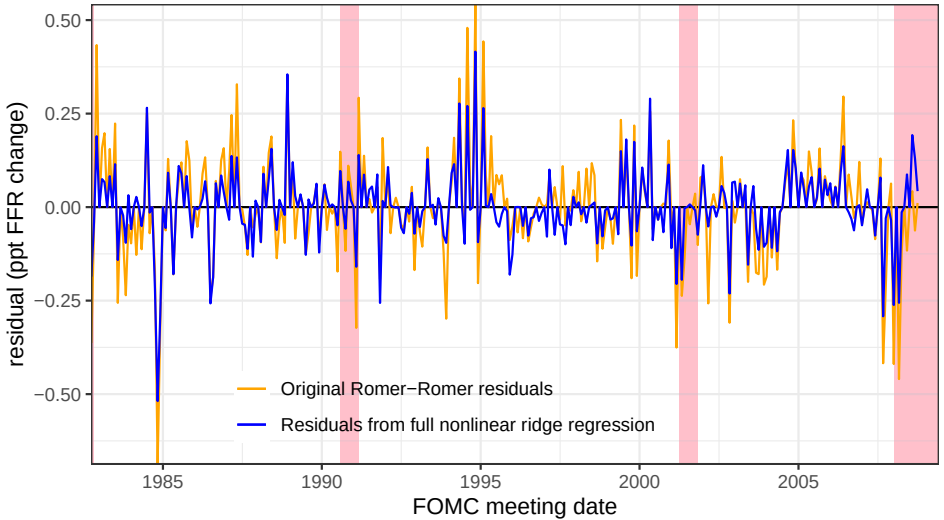
- R^2 tells us how much of the variation in Δi is explained by systematic policy
- Wider side of forecasts, human language, nonlinearities all rise R^2 robustness

WHAT EXPLAINS THE SYSTEMATIC COMPONENT?

Sentiment PC1		**Sentiment PC2**		**Numerical forecast PC1**	
economy	0.141	advanced foreign economies	-0.141	output growth (+1)	0.187
firms	0.139	merchandise	0.140	output growth (0)	0.175
economic activity	0.136	foreign economies	0.135	bus. fixed inv. growth (+2)	0.160
manufacturing activity	0.133	credit standards	-0.131	ind. prod. growth (+1)	0.160
commercial real estate	0.131	farm	0.127	output growth (+2)	0.158
manufacturing firms	0.130	cash	0.125	nominal output growth (+1)	0.153
labor market	0.125	core inflation	-0.124	housing starts (+1)	0.151
services	0.123	industrial production	0.123	housing starts (+2)	0.150
consumer confidence	0.118	trade deficit	0.121	housing starts (+3)	0.150
industries	0.117	developing countries	0.119	housing starts (0)	0.149

- Real activity variables important for sentiment and forecast PCs
- Limited role for sentiment around price and financial variables

ESTIMATED MONETARY POLICY SHOCKS



What are monetary policy shocks?

- In the paper we provide case studies for meetings with largest estimated shocks
- It turns out that these are situations in which the FOMC made decisions based on considerations not directly related to the economic outlook
  - In particular long-run credibility concerns
- Key example is November 1994 meeting, largest tighetning shock in our sample
  - Staff material suggests market had already built in a rate hike
  - Greenspan advocated a larger hike: "a mild surprise would be of significant value."

IS THERE OMITTED INFORMATION?

- We whether additional information on meeting dynamics is informative
- Allow ridge to include:
  1. Sentiment indicators from transcripts (rather than Greenbooks)
  2. Dummy variables capturing the composition of the committee

  $\Rightarrow$ expanded set of information with 1,585 variables

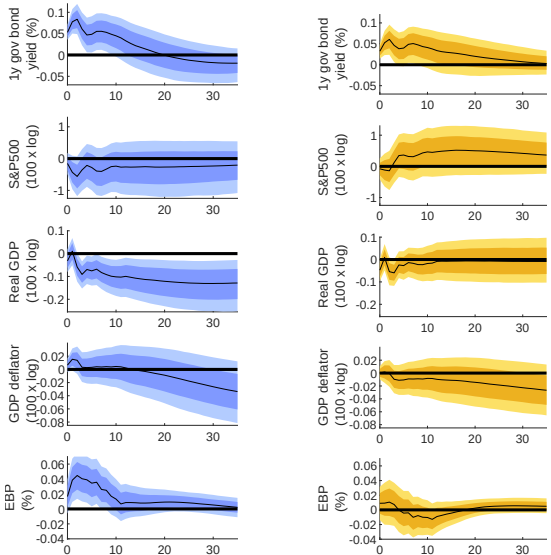
Specification	R^2
Full nonlinear Ridge	0.7505
Adding transcript sentiments and committee composition	0.7516
Difference	0.0011

THE EFFECTS OF MONETARY POLICY SHOCKS

SETTING TO ESTIMATE IRFS

- Directly follow monthly BVAR framework Jarocinski and Karadi (2020)
- Shock series is ordered first in a Choleski identification scheme (Plagborg-Moller and Wolf, 2021)
- Shock series is 1982:10 to 2008:10, but can estimate BVAR through to 2016
- System includes 1-year Treasury yield, the log of the S&P500, log real GDP, the log GDP deflator, and the excess bond premium (EBP)
- Report bands based on 16th and 84th percentiles

FULL NONLINEAR RIDGE VS. RR OLS



CONCLUSION

CONCLUSION

- Classic question in macroeconomics: what are the effects of monetary policy?
- This paper estimates monetary policy shocks by:
  - Accurately capturing the information available to the FOMC
  - Allowing for nonlinearities in the decision process
- Using natural language processing and machine learning enables us to retrieve shocks with desirable proprieties
- Monetary policy has sizeable effects on activity, inflation, asset prices, risk premia

REFERENCES

Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun (2020): "The Global Impact of Brexit Uncertainty," Tech. rep., National Bureau of Economic Research.

Jarocinski, M. and P. Karadi (2020): "Deconstructing Monetary Policy Surprises—The Role of Information Shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

Leeper, E. M., C. A. Sims, and T. Zha (1996): "What does monetary policy do?" *Brookings papers on economic activity*, 1996, 1–78.

Loughran, T. and B. McDonald (2011): "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *The Journal of Finance*, 66, 35–65.

Plagborg-Moller, M. and C. K. Wolf (2021): "Local Projections and VARs Estimate the Same Impulse Responses," *Econometrica*, 89, 955–980.

Romer, C. D. and D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94, 1055–1084.

Thornton, D. L. (2006): "When Did the FOMC Begin Targeting the Federal Funds Rate? What the Verbatim Transcripts Tell Us," *Journal of Money, Credit and Banking*, 38, 2039–2071.

APPENDIX SLIDES

COMBINING AND EXCLUDING CONCEPTS

- Using the raw list of economic concepts, we combine/exclude overlapping concepts
  - Combine singular and plural, e.g. "oil price" and "oil prices"
  - Separate mutually exclusive important concepts, e.g. keep "commercial real estate" and "residential real estate," but drop "real estate"
  - Subsume unimportant concepts if sufficiently related, e.g. drop "consumer credit" and "bank credit," but keep "credit"
  - Exclude direct mention of policy rate, since that is discussion of the action

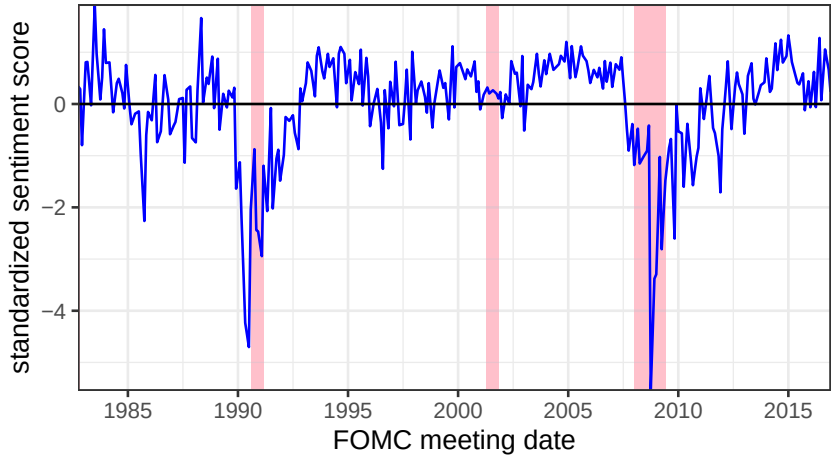
Back

3. EXAMPLES OF POSITIVE AND NEGATIVE WORDS

Positive	Negative
able	abandon
best	bad
charitable	calamities
delight	damage
easier	egregious
fantastic	fail
gain	grievances
happiest	halt
ideal	idle
leadership	jeopardize
meritorious	lack
opportunities	malfeasance
perfect	negative
...	...

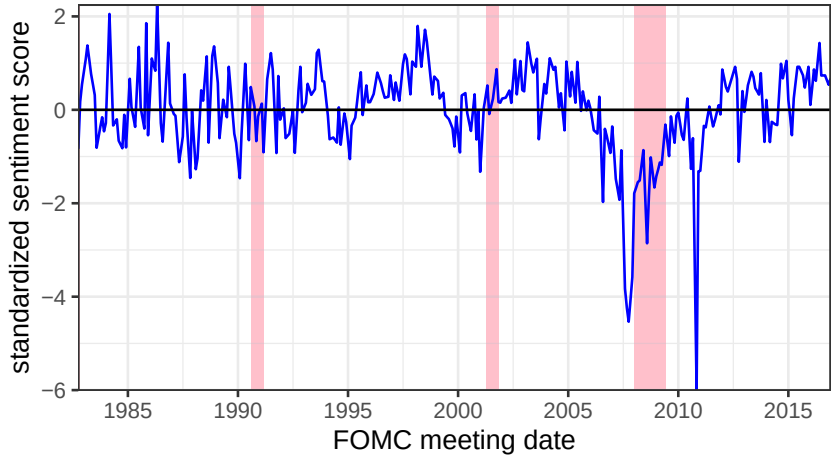
Back

CREDIT


standardized sentiment score
0
−2
−4
1985
1990
1995
2000
2005
2010
2015
FOMC meeting date

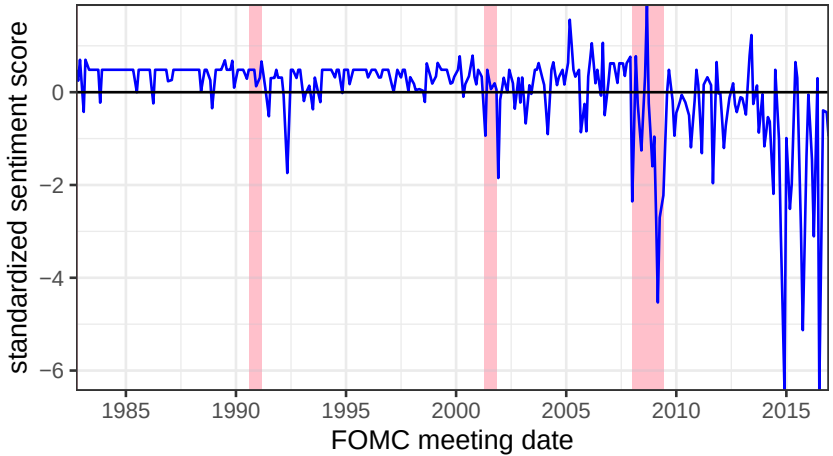
Back

MORTGAGES



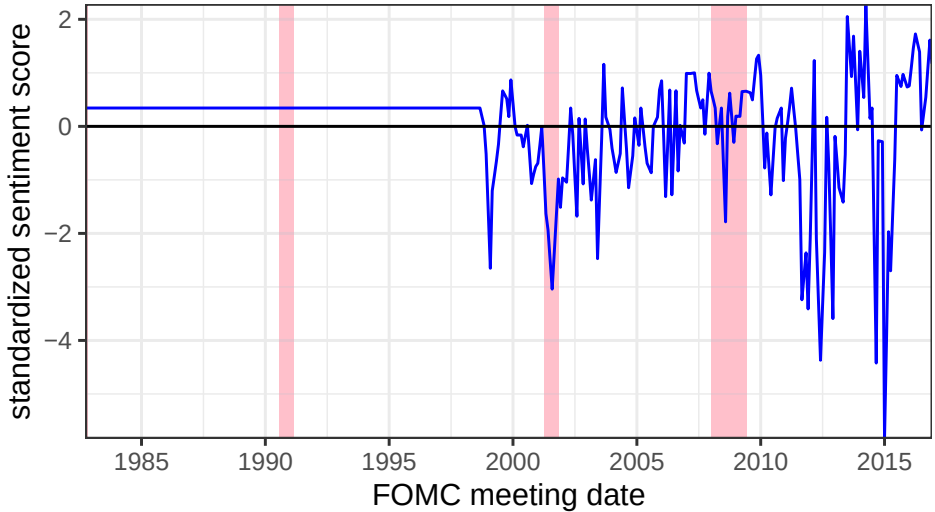
Back

INFLATION EXPECTATIONS



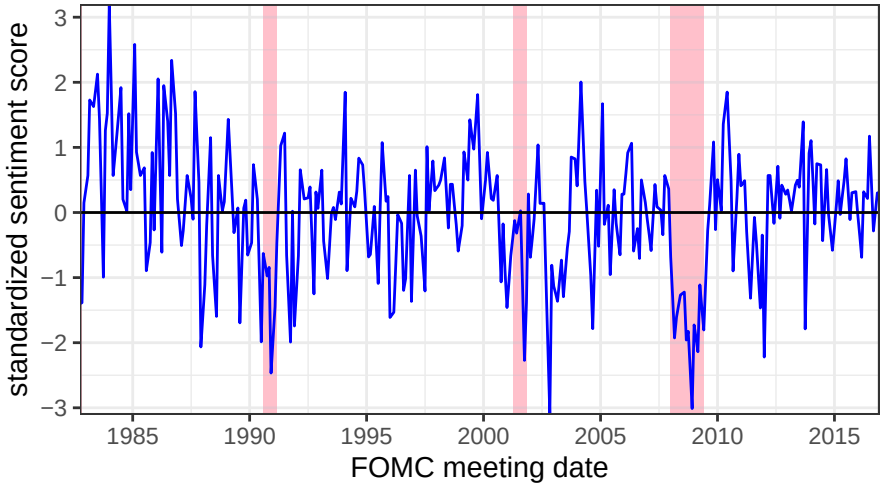
Back

EURO AREA



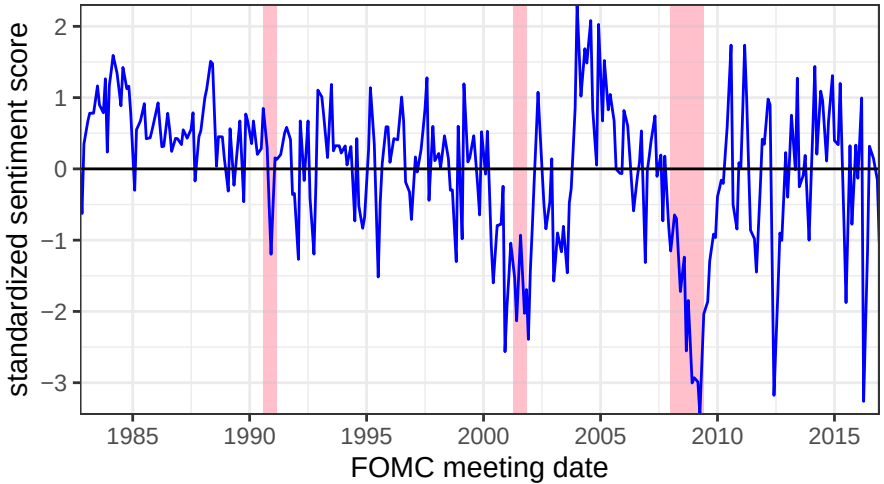
Back

CONSUMPTION



Back

LABOR MARKET



Back

k-FOLD CROSS-VALIDATION

- An optimal λ (in a predictive sense) can be found using *k-fold cross-validation*
  - Randomly divide the sample into k subsamples of equal size
  - Use each subsample fit model on the $k-1$ other subsamples
  - In each case, compute a mean-squared error (MSE) on the subsample
  - Compute an average MSE across the k MSEs
  - Find the smallest average MSE by changing λ
- We use $k = 10$

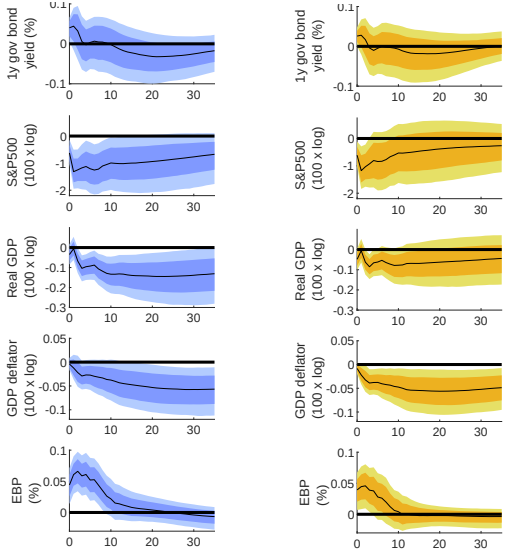
Back

R^2 ACROSS ADDITIONAL SPECIFICATIONS

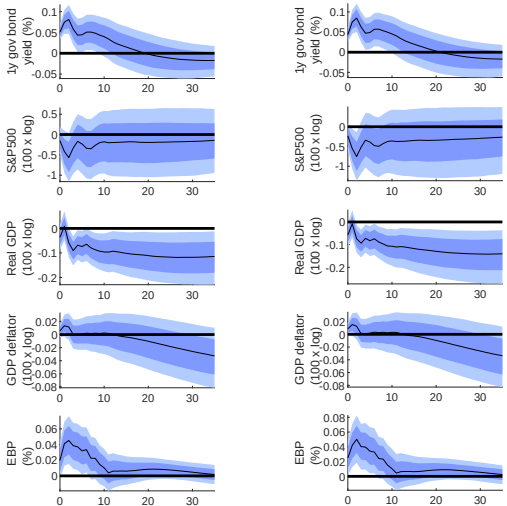
	(1) 10-word sentiment Ridge regression	(2) 5-word sentiment Ridge regression	(3) 10-word sentiment LASSO regression	(4) 5-word sentiment LASSO regression
Romer-Romer OLS	0.50	0.50	0.50	0.50
Romer-Romer ML	0.55	0.55	0.57	0.57
Full linear ML	0.65	0.66	0.55	0.63
Full nonlinear ML	0.75	0.77	0.81	0.72
Full nonlinear ML (90)	0.90	0.90	0.90	0.90

Back

FULL NONLINEAR RIDGE VS. RR OLS + JK SIGN RESTRICTIONS



INTERMEDIATE MODELS: FULL LINEAR RIDGE & RR RIDGE



RIDGE TUNING OPTION 1 VS. OPTION 2