Identifying Monetary Policy Shocks: A Natural Language Approach

Borağan Aruoba    Thomas Drechsel
University of Maryland

NBER Summer Institute 2023
MOTIVATION

- Goal: $\Delta i_t \implies Y_t$  
  Challenge: $\Delta i_t = f(\Omega_t) + \varepsilon_t$ where $Y_t \in \Omega_t$

- Romer and Romer (2004) run regression $\Delta i_t = \alpha + \beta i_{t-1} + \gamma X_t + \varepsilon_{t}^{RR}$
  - $X_t$ contains forecasts from FOMC “Greenbooks”
  - With residuals $\hat{\varepsilon}_{t}^{RR}$ construct IRFs of $Y_t$

- Key assumptions
  1. Forecasts of Fed staff good approximation of information set $\Omega_t$
  2. Linear specification good approximation of mapping $f(\cdot)$

→ this paper: both assumptions need to be revisited
Aims to revive Romer and Romer (2004) using ...

Natural language processing:
- Turn text in documents prepared for FOMC meetings into data

Machine learning:
- Include forecasts and large amount of text-based data in regression, also nonlinearly

Including the additional information is essential for clean identification
PREVIEW OF FINDINGS

1. Systematic vs. exogenous monetary policy
   ▶ Original Romer-Romer: $R^2 = 0.5$ implies 50% of $\Delta i$ are shocks
   ▶ Our approach: $R^2 = 0.94$

2. What are monetary policy shocks?
   ▶ FOMC decisions unrelated to staff’s analysis – “surprises to the staff”
   ▶ E.g. based on non-systematic long-run credibility concerns
   ▶ Correlated with high-frequency surprises in market rates

3. Get theoretically consistent IRFs in updated sample
   ▶ $i \uparrow \Rightarrow Y \downarrow \ P \downarrow \ risk\ premia \uparrow \ SP500 \downarrow$
   ▶ Not the case for original Romer-Romer because forecasts lack relevant information
METHODOLOGY
STEP 1. PROCESS RAW TEXT

- Download documents for scheduled FOMC meetings
  - Beigebook & Tealbook A (earlier: Red- & Greenbook)

- Start in 1982, when Fed began targeting FFR as policy tool (Thornton, 2006)

- End in December 2016 $\Rightarrow$ 276 FOMC meetings
  (some of subsequent analysis is pre-ZLB)

- Beigebook-only version allows us to extract shocks as FOMC meetings happen
STEP 2. IDENTIFY ECONOMIC CONCEPTS

▶ 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

▶ Select most frequently discussed economic concepts → final list amounts to 296
STEP 3. CONSTRUCT SENTIMENT

- Inspired by Hassan, Hollander, van Lent, and Tahoun (2022)
- Consider the 10 words before and after each concept’s appearance
- Each positive word gives a score of +1 and each negative word of -1
  - Classification based on enhanced version of Loughran and McDonald (2011)
- Sum up scores within meeting and scale by total number of words
EXAMPLE: SENTIMENT AROUND “ECONOMIC ACTIVITY”

![Graph showing standardized sentiment score over FOMC meeting dates from 1985 to 2015. The x-axis represents the FOMC meeting date, and the y-axis represents the standardized sentiment score. The graph includes three red vertical bars indicating significant events.]

-4 -2 0 2


FOMC meeting date

standardized sentiment score
STEP 4. RUN RIDGE REGRESSION

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

- $\widetilde{X}_t$: numerical forecasts: all variables, lags, differencing
- $Z_t$: sentiment indicators with lags
- $\Gamma(\cdot)$ captures non-linearity $\rightarrow$ implement as linear-quadratic specification

- Curse of dimensionality: up to 3,226 variables for 210 observations
- Solution: ridge regression
STEP 4. RUN RIDGE REGRESSION

- While OLS minimizes $RSS$, Ridge minimizes $RSS + \lambda \sum_{n=1}^{N} \beta_n^2$
  - Giannone, Lenza, and Primiceri (2022): dense prediction techniques tend to be preferable for economic data
  - Try alternatives, e.g. LASSO and general elastic net

- Optimally choose tuning parameter $\lambda$ based on 10-fold cross-validation
  - Maximizes ‘out-of-sample’ ability in different subsumples (folds)

- Fixed parameter ridge with large amount of information is flexible enough to also capture time-variation in the policy rule
INTERMEDIATE VALIDATION EXERCISE:
DO SENTIMENT INDICATORS PROVIDE USEFUL INFORMATION?
DO SENTIMENT INDICATORS PROVIDE USEFUL INFORMATION?


- Enough to orthogonalize FFR changes with respect to the staff’s forecasts alone ...
- ... IF forecast for variable of interest incorporates available information efficiently
- Argument relies on:

\[ \text{Greenbook forecast of } X = E[X|\Omega] \]
DO SENTIMENT INDICATORS PROVIDE USEFUL INFORMATION?

- We provide evidence that Greenbook forecast best interpreted as modal

  "I would characterize our forecasts over the years as an effort to present a meaningful, modal forecast of the most likely outcome. When we felt that there was some skewness to the probability distribution, we tried to identify it. In this instance, as we looked at the recent data, we felt that there was a greater thickness in the area of our probability distribution a little above our modal forecast."

  (Michael Prell, director of RS in FOMC meeting on July 2-3, 1996)

- We show econometrically that sentiments predict forecast errors, so $X \neq E[X|\Omega]$

  - Information from sentiments needed to clean $\Delta i_t \rightarrow$ want higher $R^2$
Do sentiment indicators provide useful information?

<table>
<thead>
<tr>
<th>Left hand side: Greenbook unemployment rate forecast errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) current quarter</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>First PC of all sentiments</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Economic activity sentiment</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

- Unemployment rate forecast errors predictable with sentiments and other variables
- Interpretation: negative activity sentiment $\Rightarrow$ positive error is consistent with negative sentiment capturing thicker upper tail
- Later: this affects whether IRFs to monetary policy shocks consistent with theory
RESULTS OF THE IDENTIFICATION PROCEDURE
$R^2$ ACROSS DIFFERENT REGRESSION MODELS

<table>
<thead>
<tr>
<th>Model Description</th>
<th>Number of regressors</th>
<th>$R^2$ with 10-word sentiment (main specification)</th>
<th>$R^2$ with 5-word sentiment (robustness)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Romer-Romer original OLS with subset of forecasts</td>
<td>19</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Ridge with extended set of forecasts</td>
<td>133</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>Ridge with all forecasts &amp; sentiments (linear)</td>
<td>429</td>
<td>0.65</td>
<td>0.66</td>
</tr>
<tr>
<td>Ridge with all forecasts &amp; sentiments (nonlinear)</td>
<td>858</td>
<td>0.75</td>
<td>0.77</td>
</tr>
<tr>
<td>Ridge with all forecasts &amp; sentiments (linear with lags)</td>
<td>1,613</td>
<td>0.87</td>
<td>0.88</td>
</tr>
<tr>
<td>Ridge with all forecasts &amp; sentiments (nonlinear with lags)</td>
<td>3,226</td>
<td>0.94</td>
<td>0.95</td>
</tr>
</tbody>
</table>

$R^2$ tells us how much of the variation in $\Delta i$ is explained by systematic policy.

Wider set of forecasts, human language, lags and nonlinearities all rise $R^2$. 

[more details on rhs]
Estimated Monetary Policy Shocks

Correlation: 0.83
WHAT ARE MONETARY POLICY SHOCKS?

- One might interpret shocks as “surprises to the staff”
- We provide case studies for meetings with largest estimated shocks
- We find that FOMC made decisions based on considerations not directly and systematically related to the economic outlook
- For **November 1994 meeting**, largest tightening shock in our sample:
  - Staff analysis suggests market had already built in a 50bp rate hike
  - Greenspan advocated a larger hike: “a mild surprise would be of significant value.”
  - The other FOMC members agree and emphasize long-run credibility
  - Increase is 75bp, we estimate a 21bp contractionary shock
OUR MEASURE VS. HIGH FREQUENCY MEASURES

- Alternative: use surprise changes in market rates around FOMC announcements
  - Gürkaynak et al. (2005), Gertler and Karadi (2015), Swanson (2021)
  - Might contain “information effect” and “Fed response to news” (Nakamura and Steinsson, 2018; Miranda-Agrippino and Ricco, 2021; Bauer and Swanson, 2023)

- Our approach orthogonalizes changes in target FFR

- Practical considerations:
  - Our shocks available over longer sample, HF measures typically start in 1990’s,
  - HF measures can be extracted from unscheduled meetings and speeches

- How do our shocks and surprises compare?
## OUR MEASURE VS. HIGH FREQUENCY MEASURES

<table>
<thead>
<tr>
<th></th>
<th>(1) Our measure</th>
<th>(2) Original Romer-Romer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation shocks with surprises</td>
<td>0.49</td>
<td>0.36</td>
</tr>
<tr>
<td>Correlation top 10 shocks with surprises</td>
<td>0.77</td>
<td>0.61</td>
</tr>
<tr>
<td>Correlation top 10 surprises with shocks</td>
<td>0.51</td>
<td>0.18</td>
</tr>
</tbody>
</table>

**Notes.** Comparison with the FFR surprises constructed by Swanson (2021). These can directly be matched to our shocks and Romer-Romer shocks for scheduled FOMC meetings. The sample period covers 1991 to 2008.

- Our shocks more strongly correlated with surprises than original Romer-Romer
- Correlation generally higher for large shocks and large surprises
- Both methods yield imperfect measures for object of interest
WHY IT MATTERS:
THE EFFECTS OF MONETARY POLICY SHOCKS
SETTING TO ESTIMATE IRFS

- Directly follow monthly BVAR of Jarocinski and Karadi (2020)
- Shock series is 1982:10 to 2008:10, but can estimate BVAR to 2016
- System includes 1-year Treasury yield, log of the S&P500, log real GDP, unemployment rate, log GDP deflator, excess bond premium (EBP)
- Report 16th - 84th and 5th-95th percentiles
- Results similar with local projections approach (Jordà, 2005)
FULL NONLINEAR RIDGE VS. RR OLS
When GB unemployment forecast error too optimistic (because mean > mode), Romer-Romer shock implies more easing than our shock.
EXPLAINING THE IMPACT ON THE RESULTING IRFS

- Suppose for simplicity $i_t$ is set based only on $\mathbb{E}(u_{t+1})$ and $\mathbb{E}(u_{t+1}) > \text{mode}(u_{t+1})$
- Predicting $i_t$ with modal forecast of $u_t$ only will imply an easing shock:

  ![Diagram showing the impact on the resulting IRFs](image)

- But this means easing shocks are estimated when unemployment goes up!
- If these instances frequent enough in the sample, the resulting IRF will be incorrect
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
▶ NLP and ML techniques enable us to retrieve shocks with desirable proprieties
▶ Monetary policy has sizeable effects on activity, inflation, asset prices, risk premia
▶ We make our estimated shocks and sentiment indicators available online!


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
## EXAMPLES OF POSITIVE AND NEGATIVE WORDS

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>adequate</td>
<td>adversely</td>
</tr>
<tr>
<td>advantage</td>
<td>aggravate</td>
</tr>
<tr>
<td>benefit</td>
<td>bad</td>
</tr>
<tr>
<td>boost</td>
<td>burdensome</td>
</tr>
<tr>
<td>confident</td>
<td>collapse</td>
</tr>
<tr>
<td>conducive</td>
<td>concerning</td>
</tr>
<tr>
<td>desirable</td>
<td>decline</td>
</tr>
<tr>
<td>diligent</td>
<td>deficient</td>
</tr>
<tr>
<td>encouraging</td>
<td>eroded</td>
</tr>
<tr>
<td>excellent</td>
<td>exacerbate</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
ALTERNATIVE SENTIMENT CONSTRUCTION

![Graph showing standardized sentiment score over FOMC meeting dates from 1985 to 2015. The graph includes two lines:
- Orange line: Extracted within sentence
- Blue line: Extracted within +/- 10 word distance
The graph highlights periods corresponding to FOMC meetings with red backgrounds.]
MORTGAGES

![Graph showing standardized sentiment score over time with FOMC meeting dates highlighted.](image-url)
INFLATION EXPECTATIONS

![Inflation Expectations Chart]

- The chart displays standardized sentiment scores over time, with FOMC meeting dates marked by vertical lines.
- The x-axis represents FOMC meeting dates from 1985 to 2015.
- The y-axis shows the standardized sentiment score, ranging from -6 to 0.
- Significant periods of economic events are highlighted in red.
LAbor market

FOMC meeting date

standardized sentiment score

### Panel (b): output forecast errors on LHS

<table>
<thead>
<tr>
<th></th>
<th>current quarter</th>
<th>1 quarter ahead</th>
<th>1 year ahead</th>
<th>2 years ahead</th>
<th>current quarter</th>
<th>1 quarter ahead</th>
<th>1 year ahead</th>
<th>2 years ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td>First PC of all sentiments</td>
<td>0.121</td>
<td>0.411</td>
<td>0.540*</td>
<td>-0.171</td>
<td>0.036</td>
<td>0.146</td>
<td>0.079</td>
<td>-0.485</td>
</tr>
<tr>
<td></td>
<td>[0.220]</td>
<td>[0.325]</td>
<td>[0.310]</td>
<td>[0.402]</td>
<td>[0.228]</td>
<td>[0.272]</td>
<td>[0.251]</td>
<td>[0.403]</td>
</tr>
<tr>
<td>Economic activity sentiment</td>
<td>0.300*</td>
<td>0.139</td>
<td>-0.252</td>
<td>-0.380</td>
<td>0.298*</td>
<td>0.131</td>
<td>-0.268</td>
<td>0.442</td>
</tr>
<tr>
<td></td>
<td>[0.167]</td>
<td>[0.276]</td>
<td>[0.340]</td>
<td>[0.750]</td>
<td>[0.163]</td>
<td>[0.299]</td>
<td>[0.374]</td>
<td>[0.717]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.005</td>
<td>0.030</td>
<td>0.049</td>
<td>0.003</td>
<td>0.000</td>
<td>0.003</td>
<td>0.001</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>206</td>
<td>204</td>
<td>198</td>
<td>54</td>
<td>206</td>
<td>204</td>
<td>198</td>
<td>54</td>
</tr>
</tbody>
</table>
## ADDITIONAL FORECAST ERROR RESULTS: INFLATION

<table>
<thead>
<tr>
<th></th>
<th>current quarter</th>
<th>1 quarter ahead</th>
<th>1 year ahead</th>
<th>2 years ahead</th>
<th>current quarter</th>
<th>1 quarter ahead</th>
<th>1 year ahead</th>
<th>2 years ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td>First PC of all sentiments</td>
<td>0.148</td>
<td>0.170</td>
<td>0.142</td>
<td>-0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.101]</td>
<td>[0.133]</td>
<td>[0.173]</td>
<td>[0.164]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economic activity sentiment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.263***</td>
<td>0.222*</td>
<td>0.236*</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.092]</td>
<td>[0.126]</td>
<td>[0.141]</td>
<td>[0.214]</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.163</td>
<td>-0.136</td>
<td>-0.267</td>
<td>0.056</td>
<td>-0.167</td>
<td>-0.140</td>
<td>-0.271</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td>[0.109]</td>
<td>[0.167]</td>
<td>[0.208]</td>
<td>[0.216]</td>
<td>[0.103]</td>
<td>[0.160]</td>
<td>[0.201]</td>
<td>[0.207]</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.029</td>
<td>0.032</td>
<td>0.017</td>
<td>0.013</td>
<td>0.081</td>
<td>0.049</td>
<td>0.041</td>
<td>0.000</td>
</tr>
<tr>
<td>Obs</td>
<td>210</td>
<td>210</td>
<td>210</td>
<td>62</td>
<td>210</td>
<td>210</td>
<td>210</td>
<td>62</td>
</tr>
</tbody>
</table>

---

Back
**WHAT EXPLAINS THE SYSTEMATIC COMPONENT?**

<table>
<thead>
<tr>
<th>Sentiment PC1</th>
<th>Sentiment PC2</th>
<th>Numerical forecast PC1</th>
</tr>
</thead>
<tbody>
<tr>
<td>economy</td>
<td>advanced foreign economies</td>
<td>-0.141</td>
</tr>
<tr>
<td>firms</td>
<td>merchandise</td>
<td>0.140</td>
</tr>
<tr>
<td>economic activity</td>
<td>foreign economies</td>
<td>0.135</td>
</tr>
<tr>
<td>manufacturing activity</td>
<td>credit standards</td>
<td>-0.131</td>
</tr>
<tr>
<td>commercial real estate</td>
<td>farm</td>
<td>0.127</td>
</tr>
<tr>
<td>manufacturing firms</td>
<td>cash</td>
<td>0.125</td>
</tr>
<tr>
<td>labor market</td>
<td>core inflation</td>
<td>-0.124</td>
</tr>
<tr>
<td>services</td>
<td>industrial production</td>
<td>0.123</td>
</tr>
<tr>
<td>consumer confidence</td>
<td>trade deficit</td>
<td>0.121</td>
</tr>
<tr>
<td>industries</td>
<td>developing countries</td>
<td>0.119</td>
</tr>
</tbody>
</table>

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

Back
OUR MEASURE VS. HIGH FREQUENCY MEASURES
LOCAL PROJECTION RESULTS

Federal Funds Rate (%)

S&P 500 Index (100 x log)

Industrial Production Index (100 x log)

Unemployment Rate (%)

Consumer Price Index (100 x log)

Excess Bond Premium (%)

Federal Funds Rate (%)

S&P 500 Index (100 x log)

Industrial Production Index (100 x log)

Unemployment Rate (%)

Consumer Price Index (100 x log)

Excess Bond Premium (%)