Keynote speech

Monetary economics and communication: new data, new tools, new and old questions

Michael McMahon, Professor of Economics,
University of Oxford

---

1 This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Monetary Economics and Communication: New Data, New Tools, New and Old Questions

Michael McMahon

University of Oxford, CEPR & Irish Fiscal Advisory Council

Oct 19, 2021

This talk represents my views and not necessarily those others including the Irish Fiscal Advisory Council, Central Bank of Ireland (co-authors), or anyone else including co-authors!
What is Central Bank Communication?

Central Bank Communication

Central bank communication broadly defined as the information that the central bank makes available about its current and future policy objectives, the current economic outlook, and the likely path for future monetary policy decisions (Blinder et al, 2008).

- **Blinder (1998):**
  
  "expectations about future central bank behavior provide the essential link between short rates and long rates."

- **Bernanke (2003):**

  "A given [monetary] policy action... can have very different effects on the economy, depending (for example) on what the private sector infers... about the information that may have induced the policymaker to act, about the policymaker’s objectives in taking the action..."
The Nature of Central Bank Transparency has Changed

Pre-1994 in US: No policy announcements, maximum opacity

- **Montagu Norman (1920-44):** 'Never apologise, never explain.’
- **Alan Greenspan (1987):** 'If I seem unduly clear to you, you must have misunderstood what I said.’

Fed since 1994: Announcements of policy & increasingly regular speeches

- **Blinder (1996):** 'Greater openness might actually improve the efficiency of monetary policy… [because] expectations about future central bank behavior provide the essential link between short rates & long rates.’

More recently (Effective Lower Bound): Forward guidance

- ⇒ Move to Open **Mouth** Operations
Why Does This New Era of Communication Matter?

What we need:

1. Answers to new questions, and new answers to old questions
2. Creation and utilisation of new data resources
3. Application and development of new methodologies

⇒ New avenues for future research

My work: 3 Related Themes

Theme I Understanding the Transmission of Monetary Policy
Theme II Monetary Policy and Expectations Management
Theme III Monetary Policy and Uncertainty
Project I: Why does CB communication move markets?

New Research with David Byrne, Robert Goodhead & Conor Parle

The Central Bank Crystal Ball: Understanding temporal information in monetary policy communication
Project I: Why does CB communication move markets?

New Research with David Byrne, Robert Goodhead & Conor Parle

The Central Bank Crystal Ball: Understanding temporal information in monetary policy communication

Key Concept

*Information Deficit*: the event must provide new, relevant information to market in order to cause them to update their beliefs.
Project I: Why does CB communication move markets?

New Research with David Byrne, Robert Goodhead & Conor Parle

*The Central Bank Crystal Ball: Understanding temporal information in monetary policy communication*

**Key Concept**

*Information Deficit*: the event must provide new, relevant information to market in order to cause them to update their beliefs.

**Key Question**

What is the nature of the central bank information deficit?
Project I: Why does CB communication move markets?

New Research with David Byrne, Robert Goodhead & Conor Parle

*The Central Bank Crystal Ball: Understanding temporal information in monetary policy communication*

**Key Concept**

*Information Deficit*: the event must provide new, relevant information to market in order to cause them to update their beliefs.

**Key Question**

What is the nature of the central bank information deficit?

**Our Approach**

Measure a new dimension of communication.
Contributions

**Contribution 1: Methodological Breakthrough**

Measuring the *temporal* dimension of text - 3rd T.

**Contribution 2: Importance of Temporal Dimension in Information Deficit**

Conjunctural context is key for policy making. Even if forward-looking aspect of policy is key for expectations, communication about context and processing of data seems to be as important as forward-looking communication.

**Contribution 3: Questions identify the deficit**

Speeches that fill the deficit are more likely to give rise to more market news.
Monetary Policy Decision-making Process

1. $\Omega^{CB}_m = g_m \left( X^{CB}_m \right)$ - Assessment Function
   
   Map data into a vector of beliefs about the state of the economy.
   
   - Macro models typically don’t focus on $g_m(.)$.
   
   - Captures two important analytical steps:
     
     - Evaluation
     - Projection
   
   - Both could be the source of $g_m(.) \neq g_{m-1}(.)$
Monetary Policy Decision-making Process

1. $\Omega_{m}^{CB} = g_{m}(X_{m}^{CB})$ - Assessment Function
   Map data into a vector of beliefs about the state of the economy.
   • Macro models typically don’t focus on $g_{m}(.)$.
   • Captures two important analytical steps:
     • Evaluation
     • Projection
   • Both could be the source of $g_{m}(.) \neq g_{m-1}(.)$

2. $i_{m} = f_{m}(\Omega_{m}^{CB})$ - Reaction Function
   Select the appropriate interest rate as a function of this state.
   • Time variation in these functions
   • No MP shock - see also McMahon and Munday (2021)
Consider the FOMC in July 2021

The Federal Reserve, the central bank of the United States, provides the nation with a safe, flexible, and stable monetary and financial system.

Federal Open Market Committee

Minutes of the Federal Open Market Committee

July 27-28, 2021

A joint meeting of the Federal Open Market Committee and the Board of Governors of the Federal Reserve System was held by videoconference on Tuesday, July 27, 2021, at 9:00 a.m. and continued on Wednesday, July 28, 2021, at 9:00 a.m.\(^1\)
Consider the FOMC in July 2021

Staff Review of the Economic Situation
The information available at the time of the July 27–28 meeting suggested that U.S. real gross domestic product (GDP) had increased in the second quarter at a faster pace than in the first quarter of the year. Indicators of labor market conditions were mixed in June, though labor demand remained strong. Consumer price inflation through May—as measured by the 12-month percentage change in the personal consumption expenditures (PCE) price index—had picked up notably, largely reflecting transitory factors.
Consider the FOMC in July 2021

Federal Open Market Committee

Staff Review of the Economic Situation

The information available at the time of the July 27–28 meeting suggested that U.S. real gross domestic product (GDP) had increased in the second quarter at a faster pace than in the first quarter of the year. Indicators of labor market conditions were mixed in June, though labor demand remained strong. Consumer price inflation through May—as measured by the 12-month percentage change in the personal consumption expenditures (PCE) price index—had picked up notably, largely reflecting transitory factors.

Available indicators suggested that growth in business fixed investment had slowed sharply in the second quarter, reflecting disruptions to motor vehicle production and aircraft deliveries and a faster rate of decline in nonresidential structures investment.
Consider the FOMC in July 2021

Federal Open Market Committee

Staff Economic Outlook

The projection for U.S. economic activity prepared by the staff for the July FOMC meeting was little changed, on balance, from the June forecast. In the second half of 2021, an easing of the surge in demand seen over the first part of the year was expected to be largely offset by a reduction in the effects of supply constraints on production, thereby allowing real GDP growth to continue at a rapid pace. For the year as a whole, therefore, real GDP was projected to post a substantial increase, with a correspondingly large decline in the unemployment rate. With the boost to spending growth from continued reductions in social distancing assumed to fade after 2021 and with a further unwinding of the effects of fiscal stimulus, GDP growth was expected to step down in 2022 and 2023. However, with monetary policy assumed to remain highly accommodative, the staff continued to anticipate that real GDP growth would outpace growth in potential output over most of this period, leading to a decline in the unemployment rate to historically low levels.

The staff's near-term outlook for inflation was revised up further in response to incoming data, but the staff continued to expect that this year's rise in inflation would prove to be transitory. The 12-month change in total and core PCE prices was well above 2 percent in May, and available data suggested that PCE price inflation would remain high in June. The staff continued to judge that the surge in demand that had resulted as the economy reopened further had combined with production bottlenecks and supply constraints to boost recent monthly inflation rates. The staff expected the 12-month change in PCE prices to move down gradually over the second part of 2021, reflecting an anticipated moderation in monthly inflation rates and the waning of base effects; even so, PCE price inflation was projected to be running well above 2 percent at the end of the year.

Over the following year, the boost to consumer prices caused by supply issues was expected to partly reverse, and import prices were expected to decelerate sharply; as a result, PCE price inflation was expected to step...
Market News

• Market Surprise: $\varepsilon^i_m = \mathbb{E}[i_m \mid \mathcal{I}^m_{mkt}] - \mathbb{E}[i_m^- \mid \mathcal{I}^m_{mkt}]$
Market News

- Market Surprise: \( \varepsilon^i_m = \mathbb{E}\left[ i_m \mid \mathcal{I}^{mkt}_m \right] - \mathbb{E}\left[ i_m- \mid \mathcal{I}^{mkt}_m- \right] \)

\[
\varepsilon_m \approx \tilde{f}_m(\Omega^m_{mkt}) - \tilde{f}_m-(\Omega^m_{mkt}) - \left[ \tilde{g}_m(X^m_{mkt}) - \tilde{g}_m-(X^m_{mkt}) \right] \tilde{f}_m-(\Omega^m_{mkt})
\]

Updated Reaction function

- Reassessment

\[- \left( X^m_{mkt} - X^m_{mkt-} \right) \tilde{g}_m'(X^m_{mkt}) \tilde{f}_m-(\Omega^m_{mkt}) \]

Effect of New Info

1. CB may choose to react to a given state of economy more or less aggressively than previously; \( f_m(.) \neq \tilde{f}_m-(.) \).
2. CB could provide more details about \( g_m(.) \neq \tilde{g}_m-(.) \)
3. CB could reveal new information not in \( X^m_{mkt} \)
Measurement: The 3 Ts of Text Analysis

- The 3 Ts
  1. *Topic* – Often measured with Latent Dirichlet Allocation (LDA)
  2. *Tone* – e.g. Dictionary Methods, VADER
  3. *Time* – Rarely captured explicitly
Measurement: The 3 Ts of Text Analysis

• The 3 Ts
  1. Topic – Often measured with Latent Dirichlet Allocation (LDA)
  2. Tone – e.g. Dictionary Methods, VADER
  3. Time – Rarely captured explicitly

Temporal dimension in our analysis

Does the 3rd T helps with understanding the nature of what drives the information deficit?

• To distinguish the temporal dimensions of communication, we use 2 approaches to time tagging:
  1. Temporal tagging via SUTime
  2. Tense tagging via TMV
Example of SUTime Output

- With **today’s** comprehensive package of monetary policy decisions, we are providing substantial monetary stimulus to ensure that financial conditions remain very favourable and support the euro area expansion, the **ongoing** build-up of domestic price pressures and, thus, the sustained convergence of inflation to our **medium-term** inflation aim.

- Let me **now** explain our assessment in greater detail, starting with the economic analysis.

- Euro area real GDP increased by 0.2%, quarter on quarter, in the **second quarter of 2019**, following a rise of 0.4% in the **previous quarter**.

- Incoming economic data and survey information continue to point to moderate but positive growth in the **third quarter of this year**.

- At the same time, the services and construction sectors show **ongoing** resilience and the euro area expansion is also supported by favourable financing conditions, further employment gains and rising wages, the mildly expansionary euro area fiscal stance and the **ongoing** albeit somewhat slower growth in global activity.
Draghi TMV Example

<table>
<thead>
<tr>
<th>Parsed Text Data</th>
<th>Verbal Complex</th>
<th>Finite</th>
<th>Tense</th>
<th>Mood</th>
<th>Voice</th>
<th>Negation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within our mandate,</td>
<td>takes</td>
<td>yes</td>
<td>present</td>
<td>indicative</td>
<td>active</td>
<td>no</td>
</tr>
<tr>
<td>the ECB is ready to do</td>
<td>is</td>
<td>yes</td>
<td>present</td>
<td>indicative</td>
<td>active</td>
<td>no</td>
</tr>
<tr>
<td>whatever it takes</td>
<td>to preserve</td>
<td>no</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>to preserve the euro.</td>
<td>to do</td>
<td>no</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>And believe me, it will be enough.</td>
<td>will be</td>
<td>yes</td>
<td>futurel</td>
<td>indicative</td>
<td>active</td>
<td>no</td>
</tr>
</tbody>
</table>
Allocations
Temporal Tagging: SUTime Measures

- We construct 2 measures of future (past) orientation using SUTime:

\[
p^{f,\text{CAT}}_s = \frac{\sum_{i=1}^{N_s^{\text{CAT}}} 1\{\text{CAT}_{i,s} = \text{Future(Past)}\}}{N_s^{\text{CAT}}},
\]

\[
p^{f,\text{NUM}}_s = \frac{\sum_{i=1}^{N_s^{\text{NUM}}} 1\{\text{NUM}_{i,s} = \text{Future(Past)}\}}{N_s^{\text{NUM}}},
\]

where \(N_s^{\text{CAT}}\) is the number of SUTime categorical references in speech \(s\), \(N_s^{\text{NUM}}\) is the number of SUTime numerical references in speech \(s\), and 1 is a dummy variable indicating when a numerical reference \(\text{NUM}_{i,s}\) or a categorical reference \(\text{CAT}_{i,s}\) is future (past).
Venn Diagram of Classifications

**Speeches**

- **Verbal Complex**
  - Temporal Reference
  - 202,077

- **Future Tense**
  - Future Temporal Reference
  - 8,250

- **Past Tense**
  - Past Temporal Reference
  - 3,751

**Statements and Answers**

- **Verbal Complex**
  - Temporal Reference
  - 31,882

- **Future Tense**
  - Future Temporal Reference
  - 2,412

- **Past Tense**
  - Past Temporal Reference
  - 953
Measuring Temporal Topics

In order to measure the topic content of the future statements, we construct topic-specific future orientation measures:

1. Identify future (past) statements using one/all of the measures:
   - SUTime Categorical
   - SUTime Numerical
   - TMV

2. Calculate the average topic share within this subset:

\[
p_{s}^{f(p),X}(T_k) = \frac{1}{N_{s}^{X=Fut(Past)}} \sum_{i=1}^{i=N_{s}^{X}} 1\{X_{i,s} = Future(Past)\} \phi_{k,m(i)}
\]

Essentially the measures are the weighted average of topic-shares from the sub-set of sentences associated with the future, according to three different metrics.
Temporal Dimensions of Monetary Policy

- FOMC Greenbooks (Tealbooks since June 2010):
  - Part 1: Summary and Outlook \(\equiv \text{Projection} = \text{Future} \& \text{Past}\)
  - Part 2: Recent Developments \(\equiv \text{Evaluation} = \text{Past}\)

(a) Past

(b) Future
Temporal Info Effects

How important is temporal information?

- Regress event news on “narrative signals” - LASSO regression
- Focus on ECB Governing Council Decisions and Press Conferences
- Use a variety of data

Key findings:
- Communication is best captured by multi-dimensional signal vectors
- Past information is about as equally as informative as future information in explaining the news.
Event Study

- Measuring beliefs via event study of asset news

\[\begin{align*}
\chi_m(t-2) & \quad \omega_m(t-2) \\
\chi_m(t) & \quad \omega_m(t) \\
\chi_m(t+1) & \quad \omega_m(t+1)
\end{align*}\]

Event Study Window

Day of Announcement

- Regress event news on “narrative signals” - LASSO regression

\[
\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{2} \sum_{i=1}^{N} (y_i - X_i \beta)^2 + \lambda \left[ \frac{1}{2} (1 - \alpha) ||\beta||^2 + \alpha ||\beta||_1 \right] \right\}
\]

- \(\alpha = 0.99\)
- Estimate \(\lambda\) by 10-fold cross-validation
- LASSO estimation via a non-parametric bootstrap - 5000 draws
- Adjusted \(R^2\) from OLS on the subset of LASSO-selected variables
Temporal Variation in ECB Statements and Answers

![Graphs showing temporal variation in ECB Statements and Answers](image)
## Expected Interest Rates

**Table:** Adjusted $R^2$ of yield curve by specification (union)

<table>
<thead>
<tr>
<th>Specification</th>
<th>OIS 1M</th>
<th>OIS 1Y</th>
<th>OIS 2Y</th>
<th>OIS 3Y</th>
<th>DE 5Y</th>
<th>DE 10Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topics Only</td>
<td>0.25</td>
<td>0.29</td>
<td>0.28</td>
<td>0.23</td>
<td>0.19</td>
<td>0.15</td>
</tr>
<tr>
<td>Topics and Future</td>
<td>0.32</td>
<td>0.36</td>
<td>0.37</td>
<td>0.31</td>
<td>0.25</td>
<td>0.19</td>
</tr>
<tr>
<td>Topics and Past</td>
<td>0.35</td>
<td>0.37</td>
<td>0.36</td>
<td>0.29</td>
<td>0.22</td>
<td>0.20</td>
</tr>
<tr>
<td>Topics, Future and Past</td>
<td>0.41</td>
<td>0.43</td>
<td>0.43</td>
<td>0.36</td>
<td>0.27</td>
<td>0.24</td>
</tr>
<tr>
<td>Topics, Future and Past*</td>
<td>0.59</td>
<td>0.62</td>
<td>0.63</td>
<td>0.59</td>
<td>0.56</td>
<td>0.47</td>
</tr>
<tr>
<td>Forecasts &amp; Revisions</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*Forecasts & Revisions column indicates whether forecasts and revisions are included in the model for each specification.*
## ABGMR factors

**Table:** Adjusted $R^2$ of ABGMR (2019) surprises by specification

<table>
<thead>
<tr>
<th>Specification</th>
<th>Target</th>
<th>Timing</th>
<th>Forward Guidance</th>
<th>Quantitative Easing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topics Only</td>
<td>0.23</td>
<td>0.28</td>
<td>0.29</td>
<td>0.21</td>
</tr>
<tr>
<td>Topics and Past</td>
<td>0.29</td>
<td>0.36</td>
<td>0.33</td>
<td>0.26</td>
</tr>
<tr>
<td>Topics and Future</td>
<td>0.37</td>
<td>0.36</td>
<td>0.34</td>
<td>0.27</td>
</tr>
<tr>
<td>Topics, Future and Past</td>
<td>0.42</td>
<td>0.42</td>
<td>0.37</td>
<td>0.32</td>
</tr>
<tr>
<td>Topics, Future and Past*</td>
<td>0.59</td>
<td>0.62</td>
<td>0.63</td>
<td>0.59</td>
</tr>
<tr>
<td>Forecasts &amp; Revisions</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Questions identifying information deficits

- It is hard to measure the information deficit
- One source is the journalists at the press conference
  - Their business model is to extract information that the market demands
- First Evidence: Questions are mix of past and future
Measuring similarity to question

- **Idea:** Typically answering questions involves talking about the same material as the question

- A measure of similarity:

  \[
  Sim_{i,j} \equiv \frac{\sum_{k=1}^{n} (dtm_{ik} \times dtm_{jk})}{\sqrt{\sum_{k=1}^{n} dtm_{ik}^2} \times \sqrt{\sum_{k=1}^{n} dtm_{jk}^2}}
  \]

  \[n = \text{number of terms in the document term matrix}\]

  \[dtm_{ik} \text{ is the } k\text{th term in the vector corresponding to document } i.\]
Similarity

Figure: Similarity Scores
**Similarity**

![Similarity Scores](image)

**Figure:** Similarity Scores
The right information constitutes news

Table: Speech-Question similarity and Information Deficit

<table>
<thead>
<tr>
<th></th>
<th>OIS 1M</th>
<th>OIS 1Y</th>
<th>OIS 3Y</th>
<th>DE 5Y</th>
<th>DE 10Y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b/se</td>
<td>b/se</td>
<td>b/se</td>
<td>b/se</td>
<td>b/se</td>
</tr>
<tr>
<td>Sim_{Sp,Q}</td>
<td>-0.23</td>
<td>13.36***</td>
<td>13.66*</td>
<td>8.79</td>
<td>6.90</td>
</tr>
<tr>
<td></td>
<td>(4.98)</td>
<td>(5.14)</td>
<td>(7.76)</td>
<td>(9.10)</td>
<td>(8.82)</td>
</tr>
<tr>
<td>Sim_{A,Q}</td>
<td>0.07</td>
<td>1.65</td>
<td>0.77</td>
<td>2.84</td>
<td>2.77</td>
</tr>
<tr>
<td></td>
<td>(1.48)</td>
<td>(1.53)</td>
<td>(2.39)</td>
<td>(2.71)</td>
<td>(2.63)</td>
</tr>
<tr>
<td>Sim_{Sp,Q} \times Sim_{A,Q}</td>
<td>-8.8</td>
<td>-40.3**</td>
<td>-39.4</td>
<td>-31.7</td>
<td>-24.6</td>
</tr>
<tr>
<td></td>
<td>(17.14)</td>
<td>(17.71)</td>
<td>(26.51)</td>
<td>(31.31)</td>
<td>(30.35)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.37</td>
<td>13.35*</td>
<td>7.10</td>
<td>14.18</td>
<td>21.32*</td>
</tr>
<tr>
<td></td>
<td>(7.03)</td>
<td>(7.26)</td>
<td>(10.77)</td>
<td>(12.85)</td>
<td>(12.45)</td>
</tr>
<tr>
<td>Speaker FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Macro controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Topics, Future and Past</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>1120.00</td>
<td>1120.00</td>
<td>1007.00</td>
<td>1122.00</td>
<td>1122.00</td>
</tr>
<tr>
<td>r2</td>
<td>0.26</td>
<td>0.41</td>
<td>0.34</td>
<td>0.22</td>
<td>0.15</td>
</tr>
<tr>
<td>Adj. R-Squared</td>
<td>0.19</td>
<td>0.35</td>
<td>0.27</td>
<td>0.15</td>
<td>0.07</td>
</tr>
</tbody>
</table>
Project 2

Cacophany of Voices

Extracting Economic Signals from Central Bank Speeches
joint with Maximilian Ahrens (University of Oxford)

- Blinder 2004:

  “A central bank that speaks with too many voices may have no voice at all.”
Our key contributions

1. We provide to the research community a novel, monetary policy shock series based on central bank speeches

2. We construct a monetary policy signal dispersion index along three key economic dimensions: GDP, CPI and unemployment
   - More frequent than FOMC meeting series
   - Opens possibilities of answering new questions regarding CB communication

3. For example, do markets form different expectations when facing a "cacophony of policy voices"? Our initial estimates suggest there might be evidence for it
Consumer price inflation has continued to move up, on balance, in recent months. The overall PCE price index rose at an annual rate of 3.9 percent during the first six months of the year, 1 percentage point faster than in the twelve months of last year. While rising energy prices have been a major source of the increase in overall consumer price inflation, the prices of core goods and services—particularly housing rents but also other core prices—have accelerated as well. Core PCE prices increased at an annual rate of 2.7 percent during the first half of this year, which is more than 1.0 percentage point higher than during 2005.

Although consumer energy prices declined 0.9 percent in June, PCE energy prices surged at an annual rate of about 25 percent over the first half of this year, up from the 17 percent increase posted in 2005. Furthermore, the downturn in energy prices in June was transitory: Spot prices for crude oil moved back up during July to a level near the peak for the year; weekly survey data on gasoline point to an increase of about 5 percent in July in the PCE price index for gasoline, a rise that would more than reverse the decline in June. Although most of the recent increase in gasoline prices reflected the higher cost of crude oil, the margin between retail gasoline prices and crude prices has moved up somewhat from an already high level despite a 35 percent decline since late June in the price of ethanol used for reformulated gasoline.

....
Text modelling

- Different text representation approaches possible (topics, word embeddings,...)

- We use a supervised topic modelling approach that learns a domain specific text representation that is optimized to predict the target variable together with other numerical covariates
  - Topic models proved to work well in relatively 'small' datasets (wide application in economics and other social sciences)
  - rSCHOLAR model (Card et al., 2018; Ahrens et al., 2021) provides ideal topic regression model for the task
  - Worth exploring different approaches for further work
Multimodal Bayesian Topic Regression

Maximilian Ahrens, Julian Ashwin, Jan-Peter Calliess and Vu Nguyen

1. The joint supervised learning approach of text representation and regression parameters allows for rigorous statistical inference when text as well as numerical features are relevant.

2. Allows us control for potential confounding factors, addressing the problem of omitted variables bias in supervised topic modelling for economic analysis.

3. Incorporating information from numerical features into the topic learning process can yield improved out-of-sample prediction performance.
Economic NLP modelling

The model estimation process is broken down into two steps:

1. Learn the mapping from central bank language to economic conditions based on Greenbook data
2. Apply the learned mapping to central bank speeches

Estimate mapping for 3 distinct economic signals:

1. GDP
2. CPI
3. Unemployment

Further work opportunities:

1. Possible to extend to additional economic dimensions
2. Possible to deploy different topic/language models to learn mapping
Example - mapping equation for CPI

Variables:

- $\Delta \pi_{4:0,m}$: change in the CPI forecast $\pi$ over the next year at FOMC meeting timestamp $m$ (target variable for CPI)
- $\theta_{\pi}$: topic mixtures for the CPI corpus. $\theta_{\{\pi, g, u\}, k}$ represents the $k^{th}$ topic feature for the respective corpus
- $\pi_{0,m-1}, g_{0,m-1}, u_{0,m-1}$: controlling for lagged variables of GDP ($g$), CPI ($\pi$) and unemployment ($u$), both in levels and in differences

Full CPI language-to-forecast mapping equation:

\[
\Delta \pi_{4:0,m} = \rho_u u_{0,m-1} + \rho_{\pi} \pi_{0,m-1} + \rho_g g_{0,m-1} + \rho_{\Delta u} \Delta u_{4:0,m-1} \\
+ \rho_{\Delta \pi} \Delta \pi_{4:0,m-1} + \rho_{\Delta g} \Delta g_{4:0,m-1} + \sum_{k=1}^{K} \omega_k \theta_{\pi,k} + \epsilon_m
\]  

(1)

$\rho$s and $\omega$s: regression weights, $\epsilon$: measurement error
Results - Estimating Implied Signals in Speeches

Figure: Out of sample implied policy signals: realised value, topic model estimation (K=20)
### Results - Estimating Implied Signals in Speeches

<table>
<thead>
<tr>
<th></th>
<th>predictive $R^2$</th>
<th>numeric</th>
<th>numeric + text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speeches - GDP signal</td>
<td>0.524</td>
<td>0.577 (0.016)</td>
<td></td>
</tr>
<tr>
<td>Speeches - CPI signal</td>
<td>0.346</td>
<td>0.575 (0.039)</td>
<td></td>
</tr>
<tr>
<td>Speeches - Unempl. signal</td>
<td>0.630</td>
<td>0.681 (0.019)</td>
<td></td>
</tr>
<tr>
<td>Greenbook - GDP training</td>
<td>0.502</td>
<td>0.766 (0.080)</td>
<td></td>
</tr>
<tr>
<td>Greenbook - CPI training</td>
<td>0.295</td>
<td>0.790 (0.147)</td>
<td></td>
</tr>
<tr>
<td>Greenbook - Unempl. training</td>
<td>0.458</td>
<td>0.657 (0.011)</td>
<td></td>
</tr>
</tbody>
</table>

**Table:** Predictive $R^2$. Models trained on Greenbook dataset, tested on speeches dataset. Best model in bold. Reported means across 50 model runs, standard errors in brackets. Numeric (OLS) has analytical solution.
Results - Estimating Speech Dispersion (CPI example)

Figure: Out of sample estimation of monetary policy signals on CPI. Top figure: signal by individual central banker speaker. Bottom figure: derived dispersion measure (grouping window: inter-FOMC-meeting periods).
Results - Estimating Speech Dispersion

Figure: Dispersion scores for GDP, CPI and unemployment compared to VIX and Economic Policy Uncertainty (EPU) index. All indices re-indexed to beginning of displayed time-series.
Estimating Dispersion Effects

• We create a single indicator:
  1. Use the interquartile range (IQR) for each of the GDP, CPI and unemployment series.
     • Simple, and protects against outliers
  2. Average these 3 dispersion series
     • Cacophony could be driven by different signals on only a subset of the indicators.
     • Therefore, alternative measure: average the two most-dispersed series
  3. This is the Dispersion Index variable

• Alternative is to create dummy indicating an intermeeting period as one of Cacophony when the Dispersion Index variable is above the median.
  • Results are similar
Market surprises are calculated using a narrow, 30-minute window around the FOMC announcement.
Effect of Cacophony on subsequent Market News

\[ |MktNews|_t = \alpha + \beta X_t + \gamma\text{Dispersion Index}_{t-1} + \epsilon_t \]

<table>
<thead>
<tr>
<th>Regressors</th>
<th>(1) Mkt News</th>
<th>(2) Mkt News</th>
<th>(3) Mkt News</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged Dispersion Index</td>
<td>0.019***</td>
<td></td>
<td>0.014***</td>
</tr>
<tr>
<td>Lagged Dispersion Index (alt)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.180</td>
<td>0.262</td>
<td>0.261</td>
</tr>
</tbody>
</table>

• Controls: NBER recession indicator, number of speeches, market volatility (VIX), policy uncertainty (BBD) and the average signal for each indicator.
Why Does This New Era of Communication Matter?

What we need:

1. Answers to new questions, and new answers to old questions
2. Creation and utilisation of new data resources
3. Application and development of new methodologies

⇒ New avenues for future research
END