Barcelona GSE
Graduate School of Economics

Hector Cano, Sam MacIntyre and Callan Windsor

# 20 Years of Central Bank Research
## What's Changed (or not) Since the Crisis?

Master in Data Science, Barcelona GSE
Master's thesis
June 2019

# 20 Years of Central Bank Research: What's Changed (or not) Since the Crisis?☆

Hector Cano, Sam MacIntyre and Callan Windsor☆

**Abstract**

We use text-mining techniques to examine every research paper published by advanced-economy central banks over the past 20 years. Our analysis points to a shift in the intellectual debate, from a relatively macroeconomic perspective toward a less aggregated view. In part, these changes seem to reflect lessons from the 2008 global financial crisis – for example, that macroeconomic models can only get you so far and that microeconomic data is useful for teasing out the causes of aggregate fluctuations. However, these changes also reflect other mutually reinforcing trends – such as an increased ability to access and analyse much larger data sets. By itself, the crisis appears to have had a sizeable and direct impact on increasing the amount of research dedicated to the banking sector, while reducing the amount of intellectual effort invested in modelling inflation.

☆We would like to thank **Hannes Muller** for his excellent supervision, in particular for helping us set our sail in the right direction. To ensure full replicability, all code used in this paper is on GitHub.

☆Email: callan.windsor at domain `barcelonagse.eu`

## 1. Introduction

It is now over 10 years since the Global Financial Crisis (GFC). Financial crises like the GFC provide central bankers with a unique opportunity to reflect on what they thought they knew about monetary economics and to re-think their modelling frameworks. Indeed, Milton Friedman's theory of monetarism was borne out of a close examination of the Great Depression (Friedman and Schwartz, 1963).

It is not surprising then that a lot of intellectual capital has been invested in looking back at the GFC. There has been an extensive search for 'lessons for central bankers' (Braude et al., 2013) and a lot of academic work dedicated to addressing shortcomings in modelling frameworks used to inform monetary policymaking (Beyer et al., 2017).

Two of the most prominent among these lessons are:

1. *models only get you so far* (Potter, 2019); and
2. *microdata is useful* (Mian and Sufi, 2010).

Reflecting on the GFC, Kohn and Sack (2018) argue that research prior to the crisis had become too narrowly focused on developing sophisticated models of the economy, and policymakers had placed too much emphasis on the predictions of these models. Central bankers had to face humbling shortcomings in their failure to foresee the crisis, despite the numerous early warning indicators that were available prior to 2008. Others suggest that central bankers had become hedgehogs: '*knowing one central truth and taking data as either affirming their views or, if not, discarding the data as unreliable, irrelevant, or uninteresting*' (Potter, 2019).[1] Consistent with this, Fligstein et al. (2014) argue that the Federal Reserve's empirical framework for making sense of the economy – macroeconomic theory – made it hard for them to connect the dots between the disparate events that ultimately led to the GFC.

The second of these 'lessons' is to pay more attention to microdata. Mian and Sufi (2010) argue that the usefulness of micro data – in particular household-level data – in teasing out the underlying causes of macroeconomic fluctuations was under-appreciated prior to the crisis. The crisis also highlighted that microeconomic issues do not wash out in aggregate and can have massive macroeconomic consequences (Simons, 2019).

Against this background, the aim of this paper is to see if the rhetoric regarding these two 'lessons' matches the reality. We examine trends in the intellectual debate occurring

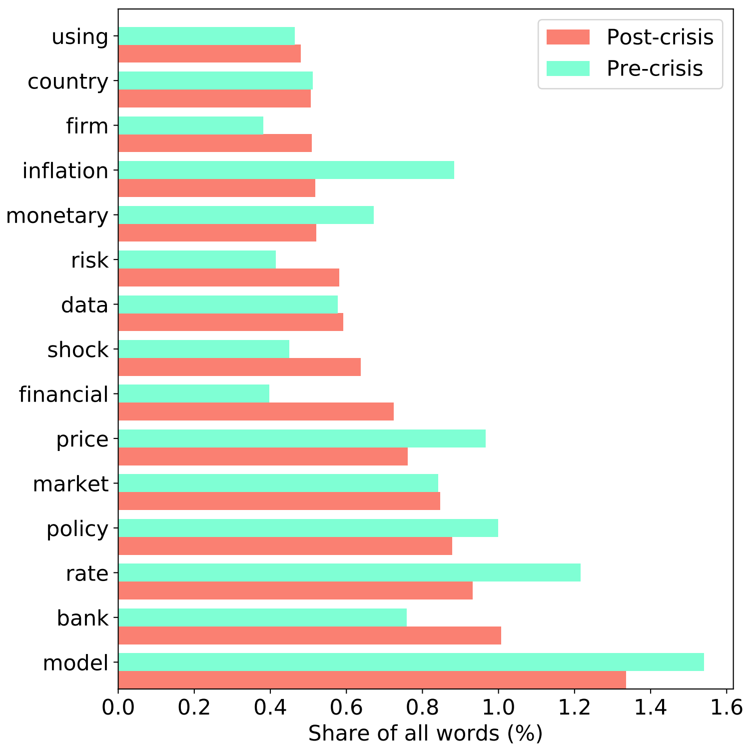
[1] The term *hedgehog* in this context was introduced by Tetlock and Gardner (2015). They define a hedgehog as a specialized and rigid forecaster, while the *fox* represents a broader breed of thinker. The hedgehog tends to have some grand theory that informs all they do and say, while the fox is a generalist.

within central banks over the past 20 years by analysing the research papers they publish. We also attempt to see if the intellectual debate shifted following the crisis in a way that is consistent with the two lessons outlined above. *Because, if the research hasn't changed, can it really be said that policy makers have learnt anything?*

In particular, we apply techniques from computational linguistics to examine a census of research papers written by inflation targeting central banks, covering 10 years before and 10 years after the crisis. Our goal is to try and focus on what has changed intellectually, rather than what has changed on the policy front. Glancing at the most commonly used words in central banking research over the past two decades, it is of little surprise that terms such as 'financial', 'risk', 'bank' and 'shock' have increased in frequency since the GFC (Figure 1).

**Figure 1: Word Frequencies in Central Bank Research**

Advanced economy central banks: 2000–2019




Note: Word count calculated after tokenizing, removing stop words and lemmatizing
Sources: authors' calculations; IDEAS

To dig deeper and find patterns in the text that are consistent with the intellectual shifts described above, we employ three empirical techniques:

1. First, we use **topic models** – Latent Dirichlet Allocation (LDA) (Blei et al., 2003) and dynamic topic models (Blei and Lafferty, 2006) – to examine if there has been a shift away from modelling and towards a greater focus on micro data. These models are useful in discovering topics from a large and unstructured collection of documents.

2. In a similar vein, next we use **latent semantic analysis** (LSA) to examine if central bank research has moved away from a narrow focus on modelling (our approach here is similar to Iaria et al. (2018)). In particular, we measure the similarity of central bank research to frontier research dedicated to macroeconomic modelling. To characterise frontier research, we build an index using research papers from two of the highest rated macro journals: *The Journal of Macroeconomics* and *The B.E. Journal of Macroeconomics*.

3. Finally, we attempt to distinguish between longer-term changes in the intellectual debate versus something directly attributable to the financial crisis. To do this, the results from the topic modelling exercise are used in a regression setting similar to **regression discontinuity design**. This allows us to control for changes in the intellectual debate that were occurring before the onset of the crisis.

While our results are mostly descriptive in nature and should not be over-interpreted, the weight of the evidence suggests that the rhetoric appears to match the reality. The results from the topic modelling exercise indicate that there has been a shift toward a greater focus on microdata alongside a reduction in the emphasis on modelling over the past 20 years. Similarly, we find that central bank research has become more 'distanced' from frontier research dedicated to macro modelling since the crisis – as measured by paper similarity scores in a **difference-in-difference** setting. However, only a handful of topics appear to have been directly affected by the crisis. There was a notable increase in the weight of research dedicated to the banking sector and a decrease in research dedicated to inflation modelling.

The rest of this paper is organized as follows: Section 2 outlines the web scraping tools used to collect the research abstracts we analyse. Section 3 outlines the topic modelling methodology; describes the identified topic landscape; and examines how topics have evolved over time. Similarly, Section 4 reviews the LSA methodology and presents the results from this exercise. Section 5 then uses the results from the LDA topic model in a regression discontinuity setting to try and tease out the direct impact of the crisis on topic weights. Finally, the conclusions are provided in Section 6, along with limitations.

## 2. Data

### 2.1. Collection

We scrape the IDEAS economic research repository to gather the research abstracts analysed in this paper. IDEAS is the largest bibliographic database dedicated to economics and available freely on the internet. In particular, we collect every research abstract published by nine advanced economy inflation-targeting central banks over the period 2000–2019 (Table 1).

Likewise, to compare central bank paper abstracts to academic research abstracts (with similarity evaluated by LSA) we scrape the IDEAS pages of the two highest rated macroeconomic journals: *The Journal of Macroeconomics* and *The B.E. Journal of Macroeconomics*. These journals publish significant theoretical and empirical research spanning the entire range of macroeconomics and monetary economics.

**Table 1: Web Scraping for Research: 2000–2019**

| Central bank | No. of abstracts |
|---|---|
| European Central Bank Working Paper Series | 2285 |
| Federal Reserve | 1498 |
| Bank of Canada Staff Working Papers | 895 |
| Bank of England Working Papers | 687 |
| Norges Working Paper | 335 |
| Bank of Japan Working Paper Series | 281 |
| Sveriges Riksbank Working Paper Series | 262 |
| RBA Research Discussion Papers | 210 |
| Reserve Bank of New Zealand Discussion Paper Series | 184 |
| Total | 6637 |
| Academic journals | |
| The Journal of Macroeconomics | 1400 |
| The B.E. Journal of Macroeconomics | 581 |
| Total | 1981 |

Note: The python code used to perform the scraping exercise is on GitHub.
Sources: authors' calculations; IDEAS

We focus only on advanced economy inflation-targeting central banks to control for fundamental factors that shape research agendas. The central banks of emerging market economies tend to focus their research efforts on a very different set of issues. Likewise, lessons from the GFC are arguably more relevant for advanced economies, just as lessons from the Asian Financial Crisis or the Latin American Debt crisis were more relevant for emerging market economies.

A natural question to ask is why we do not focus exclusively on the US Federal Reserve, given the crisis originated in the United States. We choose to focus on other advanced economy central banks because lessons learned from the crisis are very relevant for these economies too. This notwithstanding, results are presented separately for the Federal Reserve where relevant.

### *2.2. Text preprocessing*

After collecting the abstracts, text pre-processing was required before we made use of topic models to examine the evolution of topics over time as well as LSA to examine the similarity of research abstracts.

For every abstract, all punctuation and numbers were removed and all characters were made lower-case. We then eliminated all stop-words, which are typically 'function' words. These words have very little substantive meaning and primarily denote grammatical relationships between 'content' words, such as prepositions (of, by, from), conjunctions (and, but, till, as) and in/definite articles (a, an, the). In addition, we remove 'domain-specific' stop words, which do not add any semantic value. Words that appear in most economic articles were removed, such as 'study', 'paper', 'find', 'effect', 'discuss', 'suggests','implies' and 'indicates'.

'Lemmatization' was then conducted to find the lemma – or dictionary form – of each word in order to further reduce dimensionality (i.e. the total number of words), without losing generality. We take the lemma of each word rather than its stem because it better preserves both meaning and part-of-speech information. For example, the lemma of the word *productivity* is productivity, whereas its stem is *product* – a semantically very different word. Since the identification of topics relies on the subjective interpretation of word distributions, lemmatization is more appropriate than stemming in our case. In all, preprocessing reduces the number of unique words in our corpus from 15,159 words to 13,251 unique words (Table 2).

**Table 2: Text Preprocessing**

| | Raw text | Stopwords | Domain stopwords | Lemmatization |
|---|---|---|---|---|
| Words | 923,555 | 577,729 | 546,483 | 546,483 |
| Unique words | 15,159 | 15,039 | 15,000 | 13,251 |

Notes: Our raw tokenised text contains 923,555 words, of which 15,159 are unique. After removing standard stopwords; our domain-specific list of stop words; and lemmatizing there are 546,483 words, of which 13,251 are unique
Sources: authors' calculations; IDEAS

The final preprocessing step is to transform the research abstracts into a $D \times V$ document-term matrix – or a bag-of-words. In a bag-of-words, each document is represented in a vector of an unordered collection of words – for a total of $V$ words in a corpus, each abstract becomes a $V$-dimensional vector .

For the LSA, the individual word counts in the $D \times V$ matrix are reweighted by their term frequency inverse document frequency (tf-idf) using Equation 1 below. Here, if the frequency of term $i$ in document $j$ is high then the weight of that term will be high; however, if the term appears in a lot of documents, the weight will decline. This transformation essentially decreases the relative importance of words that carry little information, but appear in many documents.

$$\text{weight}_{i,j} = \text{frequency}_{i,j} * \log_2 \frac{\text{Number of documents}}{\text{document freq}_i} \tag{1}$$

## 3. Topic models

To examine the themes in central banking research we use the very popular **LDA** algorithm developed by Blei et al. (2003). This allows us to compare the weight of various topics found in central banking research and how the relative importance of these topics have changed over time. We also use a **dynamic topic model**, which respects the dates the research papers were published (Blei and Lafferty, 2006). This lets us to look within topics over time and examine how the topic itself has evolved. Below we briefly describe the intuition behind these topic models; a more in-depth discussion of both topic models is provided in Appendix A.

Starting first with LDA, we begin by passing the algorithm two **two inputs** . The first is the preprocessed corpus of all central bank research abstracts published since 2000. The second is the number of topics represented in each research paper abstract. Topics are subjectively labelled with reference to their fixed distribution over words. For example, we would expect the topic labelled *macro modelling* to assign a high probability to the word *forecast*. We use a 10-topic model, which ensures that the identified topics are easily interpretable.[2] However, this comes at the cost of model precision, with a larger number of topics improving the performance of the model (see Chang et al. (2009) for a summary of this trade-off).

[2] In a similar context, Hansen et al. (2017) use a 15-topic model to examine the effect of central bank communication on both market and real economic variables.

Upon completion of LDA inference, there are **two outputs**. The first is a $D \times K$ per-document topic proportion matrix $\theta$. In our case, the dimension of this matrix is around $6600 \times 10$. The second is a $V \times K$ per-topic word distribution matrix $\Phi$, with a dimension of around $13500 \times 10$. An example of these outputs is provided in Table 3.

**Table 3: LDA Output Examples**

**Per-document topic proportions** $(\theta_d)$

| | $T_1$ | $T_2$ | ... | $T_K$ | Sum |
|---|---|---|---|---|---|
| $Abstract_1$ | 0.2 | 0.5 | ... | 0.1 | 1 |
| $Abstract_2$ | 0.5 | 0.2 | ... | 0.1 | 1 |
| ... | ... | ... | ... | ... | 1 |
| $Abstract_D$ | 0.9 | 0.0 | ... | 0.0 | 1 |

**Per-topic word distributions** $(\phi_k)$

| | $T_1$ | $T_2$ | ... | $T_K$ |
|---|---|---|---|---|
| $word_1$ | 0.01 | 0.02 | ... | 0.02 |
| $word_2$ | 0.02 | 0.02 | ... | 0.01 |
| ... | ... | ... | ... | ... |
| $word_V$ | 0.04 | 0.01 | ... | 0.01 |
| Sum | 1 | 1 | 1 | 1 |

Under the assumptions of the simple LDA model, the temporal order of documents does not matter (that is, the documents are exchangeable). This assumption may be unrealistic when analysing long-running collections that span several decades, such as our collection of central bank research abstracts. As a result, we also employ a *dynamic topic model*, where we assume that the distribution of words that define a topic evolves over time. This approach lets us look at the change *within* topics and track how it has changed.

To do this, the corpus is divided into 20 yearly time slices, with the assumption that within each slice documents are exchangeable. The topic distributions over words are then allowed to evolve from slice to slice. The intuitive idea is given by Equation 2.

$$\beta_{t,k} | \beta_{t-1,k} \sim \mathcal{N}\left(\beta_{t-1,k}, \sigma^2 I\right) \tag{2}$$

Here, $\beta_{t,k}$ denotes the $k^{th}$ topic at time $t$, given that same topic in the previous year. This is distributed as a normal, whose mean is the topic in the previous year along with a covariance $\sigma^2$, which represents how much the topic can move from year to year (see Appendix A).

Equation 2 is then mapped to a probability distribution over words. The hyper-parameter $\sigma^2$ is important as it governs how fast or slow topics evolve over time. If $\sigma^2$ is too large then the model will be able to 'explain' new topics by completely changing the set of words over which that topic is defined. In practice, we set $\sigma^2$ to the default value of 0.005 as suggested by (Blei and Lafferty, 2006) in the C++ code accompanying their paper.

### *3.1. The topic landscape*

The extracted and labelled 10 topics from central banking research are presented as word-clouds in Figure 2. As indicated by the colour scheme, the topics neatly resemble the key responsibilities of a typical inflation-targeting central bank, namely:

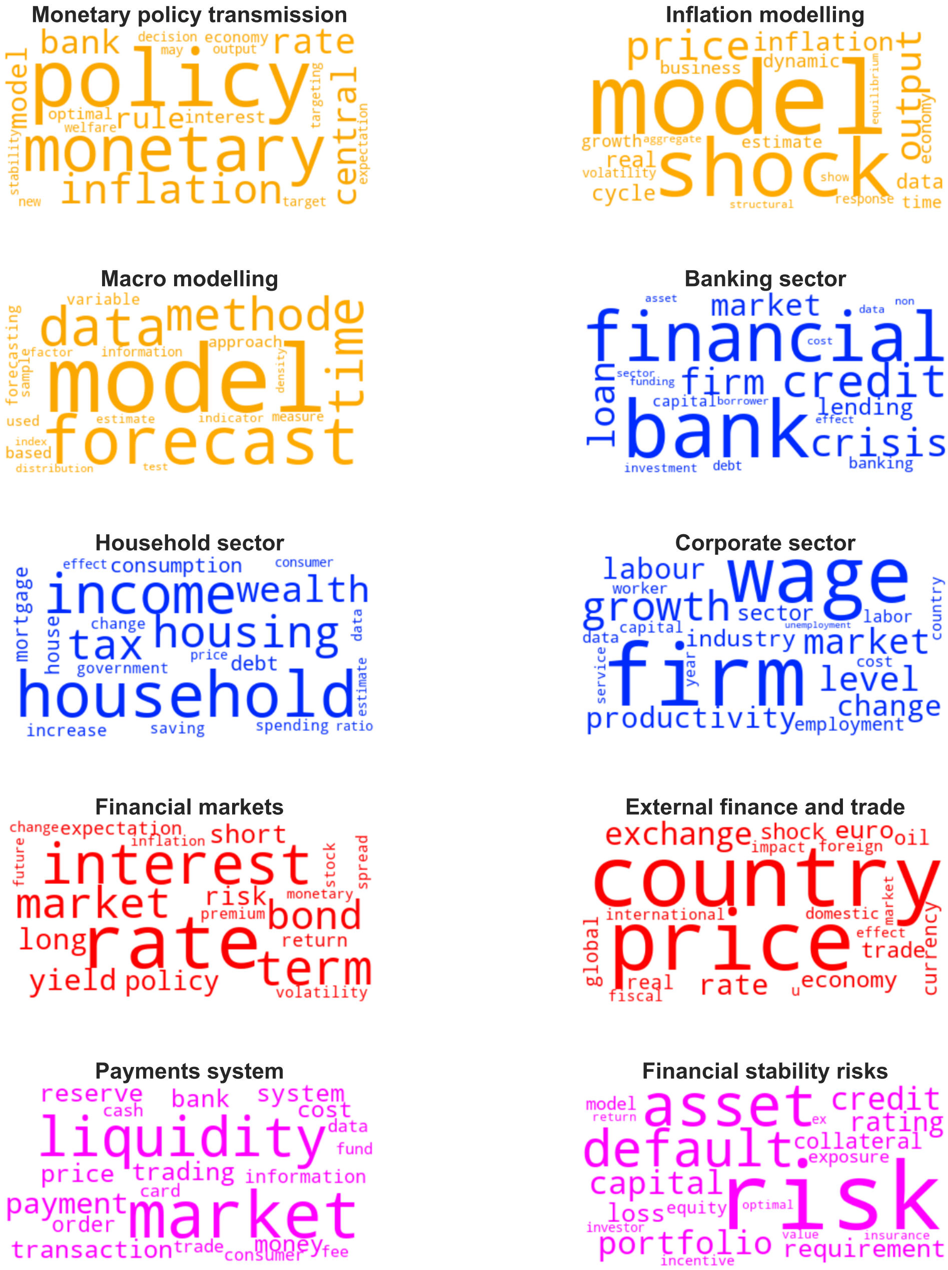
- **modelling:** *policy transmission*, *inflation modelling*, *macro modelling*;
- **economic analysis**: *banking sector*, *household sector*, *business sector*;
- **financial markets**: *financial markets*, *external finance and trade*;
- **the payments system**; and
- **financial stability**.

The topics also neatly cluster into intuitive groups. Figure 3 shows that the modelling topics (*inflation modelling*, *macro modelling* and *policy transmission*) are closely related to each other. Likewise, topics focusing on individual sectors of the economy (households, corporates and banks) cluster into a neighbourhood. Unsurprisingly, the topic *financial stability risks* is also tightly connected to the *banking sector* topic.

To get a sense of the importance of each topic, Table 4 sorts topics according to their proportions in the whole collection of abstracts (i.e. by taking the average of each topic in the topic proportion matrix $\theta$). The number of abstracts in which each topic has the highest proportion is also provided, together with the six most frequent words for a given topic. Appendix B.8 plots the word count and importance of keywords for each topic.

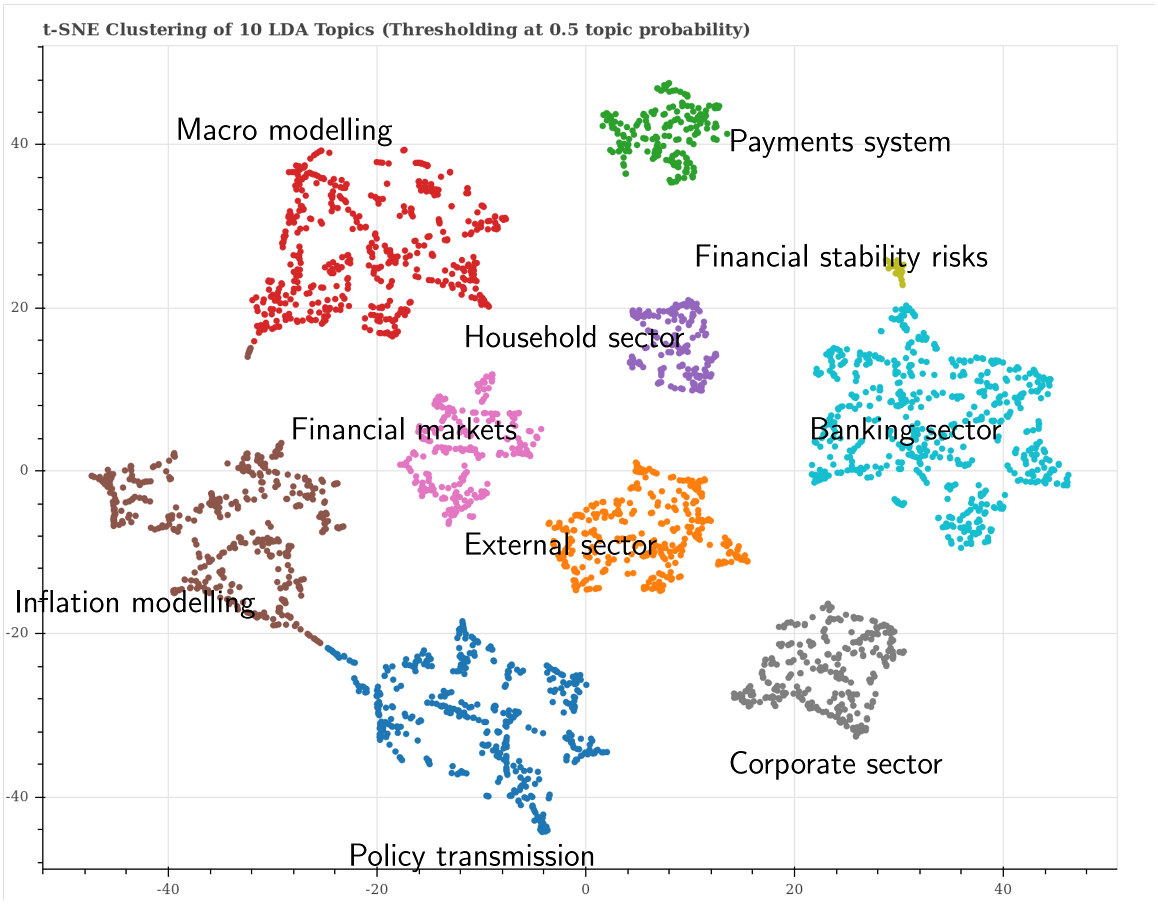
Over the period 2000–2019, the *banking sector* has had the highest weight in central banking research. This is due to a significant increase in the post-crisis period, as explained in the next section. Unsurprisingly, papers dedicated to *policy transmission*, *inflation modelling* and *macro modelling* have also historically been important themes in central banking research. Receiving less attention have been topics related to the individual sectors of the economy (i.e. households and corporates).

**Figure 2: Topics Covering Central Banking Research**




Sources: authors' calculations; IDEAS

**Figure 3: The Clustered Topics of Central Bank Research**




Notes: This Figure uses the t-distributed stochastic neighbour embedding (t-SNE; (van der Maaten and Hinton, 2008)) – a popular dimensionality reduction algorithm – to reduce the topic-proportions for each abstract into a two-dimensional space, with each dot representing a central bank research abstract

Sources: authors' calculations; IDEAS

**Table 4: Topics of Central Banking Research: 2000–2019**

| Top 6 topic words | No. of abstracts | Proportion (%) | Topic label |
|---|---|---|---|
| bank financial credit crisis loan firm | 1183 | 0.139 | Banking sector |
| model shock price output inflation real | 895 | 0.136 | Inflation modelling |
| model forecast data time method approach | 864 | 0.117 | Macro modelling |
| policy monetary inflation rate central bank | 814 | 0.116 | Monetary policy transmission |
| country price exchange rate euro economy | 764 | 0.107 | International finance and trade |
| firm wage growth market level productivity | 590 | 0.084 | Corporate sector |
| rate interest term market bond yield | 577 | 0.081 | Financial markets |
| market liquidity payment transaction price reserve | 412 | 0.054 | Payments system |
| household income housing tax wealth consumption | 428 | 0.052 | Household sector |
| risk asset default portfolio capital credit | 110 | 0.024 | Financial stability risks |

Notes: The column *no.of abstracts* indicates the number of times a given topic had the highest weight in a central bank research abstract; the column *proportion* indicates the mean topic proportion across all abstracts
Sources: authors' calculations; IDEAS

### 3.2. *LDA topic trends: how fixed topics have evolved over time*

Against this landscape, we now examine how the intellectual debate within central banks has evolved over time. In particular, we examine if the topics have shifted in a way that is broadly consistent with the two so-called 'lessons learned' since the crisis, namely: (1) that *models only get you so far*; and (ii) that *micro data is useful*.

To examine topic trends over time using the simple LDA model, we collapse the topic-proportion matrix ($\theta$) by taking the mean of the topic proportions for all abstracts by year. This approach to examining the changing weight of fixed topics over time has been popular in the literature (see, for example, Lee and Kang (2018)).

The time-varying topic proportions are then regressed on a linear time trend to identify those topics with significantly increasing or decreasing topic weights over the 20-year period. Topics whose regression slopes are positive (negative) at a significance level of 1% are plotted in Figure 3.

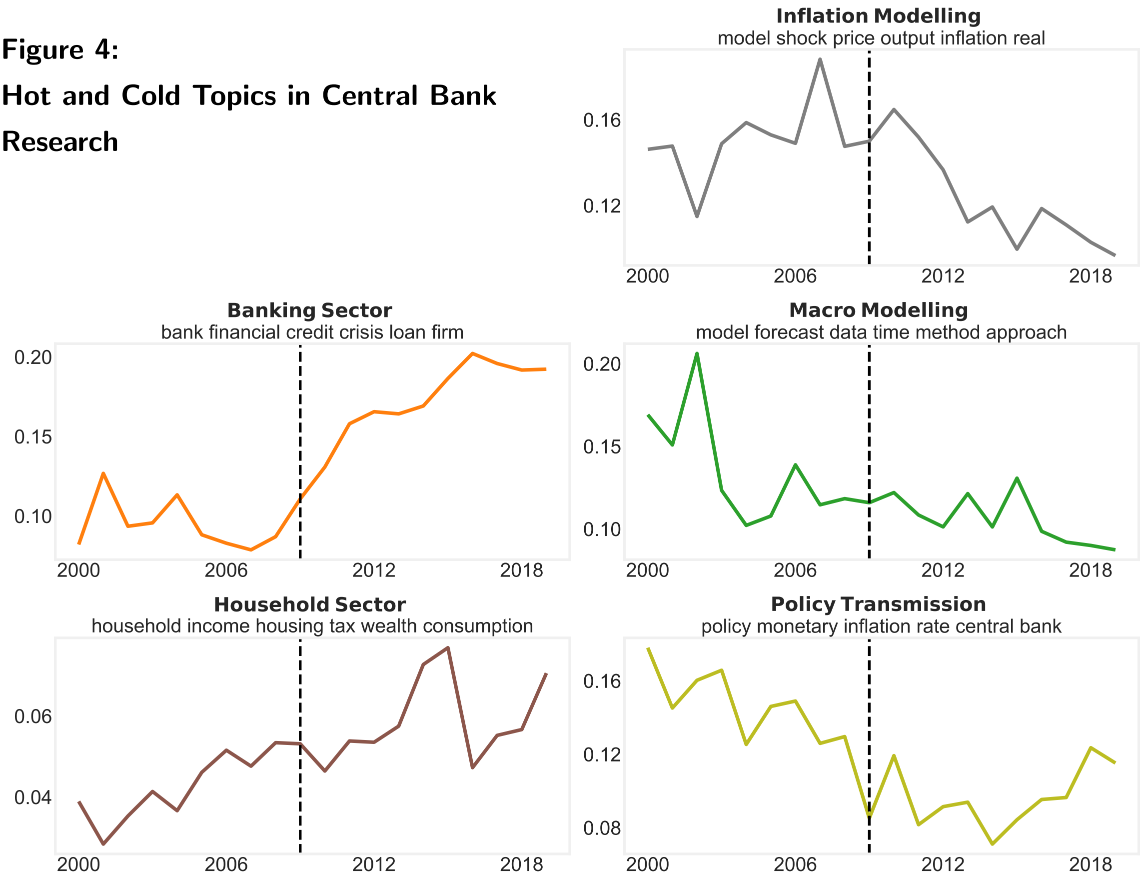
We observe that research dedicated to the *banking sector* has significantly increased since the 2000s, especially so since the crisis. Research on the *household sector* topic has also increased, though the impact of the crisis is difficult to discern. The topic proportion exhibits a significant increase both before and after the crisis. On the other hand, research output falling under the broad umbrella of modelling (*policy transmission*, *inflation modelling* and *macro modelling*) has significantly declined.

At face value, these trends are consistent with central banks looking beyond their macro models and instead adopting a less aggregated perspective. Again, however, it is difficult to identify a noticeable impact of the financial crisis on these trends.

Overall, it is likely that these trends reflect lessons from the crisis as well as a number of other mutually reinforcing factors. For example, there has been an increasing ability to create, access and analyse much larger data sets over our 20-year sample period. As more researchers have seen the potential in these datasets, there has been an rise in associated research output.

The fall in research dedicated to *monetary policy transmission* and *inflation modelling* could also be a sign of learning. Simons (2019) suggests that as questions about the monetary policy framework have been adequately answered there are less questions to be asked and so less research produced. That is, despite the crisis, circumstances have not changed in such a way as to reveal a weakness in the monetary policy framework.

**Figure 4:**
**Hot and Cold Topics in Central Bank Research**




Notes: For significant time trends at the 1 per cent level; for each topic, the subtitle lists the top six words by weight
Sources: authors' calculations; IDEAS

To examine changes in the mean topic weight for each topic over the pre- and post-crisis periods, we estimate Equation 3 below for all central banks and for the US Fed in isolation:

$$T_{i,a,t} = \sum_{t=1}^{T=2} \sum_{i=1}^{I=10} \lambda_{i,t} D_{i,t} + \varepsilon_{i,a,t}. \tag{3}$$

Here, $T_{i,a,t}$ denotes topic $i's$ proportion in abstract $a$ in time period $t$ (pre or post crisis); $D_{it}$ is a dummy variable for topic $i$ in time period $t$; and the coefficient on $\lambda_{it}$ captures the mean topic proportion for topic $i$ in the pre and post crisis period. With this simple regression, for each topic we can examine if the mean topic proportion shifted significantly in the post-crisis period.

For all central banks, there has been a significant decline in *inflation modelling*, *macro modelling* and *monetary policy transmission* since the crisis (Table 5). On the other hand, topics related to the *banking sector* and the *household sector* have increased significantly, while the focus on the *corporate sector* has remained unchanged.

**Table 5: Pre- and Post-Crisis Topic Proportions**

| | All central banks | US Fed |
|---|---|---|
| *Pre-crisis topic proportion* | | |
| Inflation modelling | 0.15*** | 0.11*** |
| Macro modelling | 0.13*** | 0.11*** |
| Policy transmission | 0.14*** | 0.11*** |
| Financial mkts | 0.08*** | 0.12*** |
| External finance | 0.11*** | 0.03*** |
| Banking sector | 0.09*** | 0.12*** |
| Household sector | 0.04*** | 0.10*** |
| Corporate sector | 0.09*** | 0.13*** |
| Payments system | 0.05*** | 0.06*** |
| Financial stability | 0.02*** | 0.03*** |
| *Change in topic proportion* | | |
| post crisis× Inflation modelling | -0.03*** | 0.00 |
| post crisis× Macro modelling | -0.02*** | -0.01 |
| post crisis× Policy transmission | -0.05*** | -0.02** |
| post crisis× Financial mkts | 0.00 | -0.02* |
| post crisis× External finance | -0.01** | 0.01 |
| post crisis× Banking sector | 0.08*** | 0.06*** |
| post crisis× Household sector | 0.02*** | 0.00 |
| post crisis× Corporate sector | -0.00 | -0.03*** |
| post crisis× Payments system | 0.01*** | 0.01* |
| post crisis× Financial stability | 0.01*** | 0.00 |
| Observations | 66370 | 14980 |

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ using HAC robust standard errors;
the number of observations is equal to the number of abstracts times the number of topics $(6637 \times 10)$
Sources: authors' calculations'; IDEAS

Post-crisis shifts in the intellectual debate are less evident when looking at the US Fed in isolation. There is no evidence of a significant change in the weight given to either *macro* or *inflation modelling*. Moreover, while there has been a significant decline in the topic weight on *monetary policy transmission*, the fall is less pronounced relative to other central banks. Of the individual sectors, only the banking sector has increased in weight in the post-crisis period.

These results seem to suggest that the intellectual debate within the Fed suffers from more inertia relative to other central banks, despite the crisis originating in the United States. However, an alternative (and perhaps more convincing) explanation is that the Federal Reserve has been ahead of the learning curve. Relative to other central banks, the Fed's pre-crisis research was less heavily weighted toward modelling topics and instead placed more emphasis on research about the individual sectors of the economy. This is neatly illustrated in Appendix C, which plots the relative topic strength for each central bank.

### *3.3. Dynamic topic trends: how topics have evolved over time*

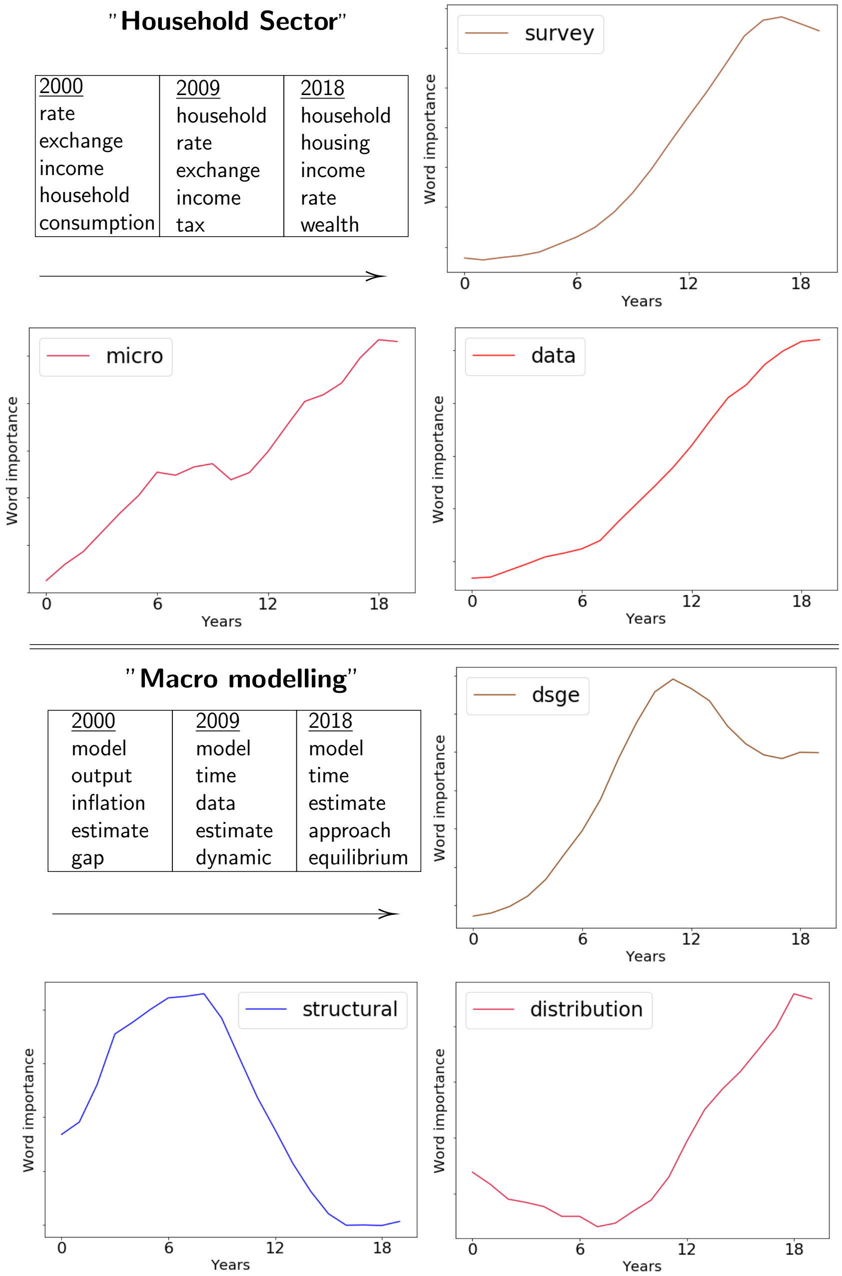
To dig deeper and examine how topics themselves have evolved from year-to-year, we now turn to the results from the dynamic topic model. Here, we focus on the *household sector* and *macro modelling* topics, as these are most closely related to the changes in the intellectual debate we expect to observe.

Figure 5 illustrates the top five words from the topic model before (2000), immediately after (2009) and following the crisis (2018). It also shows estimated word importances as a function of each year for three words from each topic that describe important changes in the topic definition over time. Appendix D shows the top 30 words for each topic and selected time slice (2000, 2009, 2018) along with the associated word importances.

Interestingly, trends in word usage for the *household sector* reveal that the topic has evolved over time to incorporate a more micro perspective. The words *survey*, *data*, *micro*, *distribution* and *inequality* have become more important in defining the *household sector* topic over time. In addition, looking at the five words with the highest weight in the topic definition, it appears that the topic has evolved from a focus on the impact of external-sector shocks to a focus on the effect of domestic factors. For example, the words *exchange* and *currency* have declined noticeably since 2000 and instead words like *mortgage*, *debt* and *housing* have increased in importance.

Trends in the word usage for the *macro modelling* topic seem to indicate an evolution away from large-scale structural modelling. For example, since the crisis, the term *DSGE* (the acronym for Dynamic Stochastic General Equilibrium models) has played a declining role in defining the topic. Likewise, the term *structural* has declined in importance over time. By contrast, terms such as *distribution*, *micro*, *sector*, *financial* and *friction* have increased. These trends speak to a shift from a relatively macroeconomic perspective toward incorporating financial frictions and micro-foundations into macro modelling and theory. Some of these changes were occurring before the crisis, but it is possible that the crisis brought about a deeper realisation that traditional large-scale macro models were not providing the correct signals in the lead up to the crisis.

**Figure 5: Dynamic Topic Evolution**



## 4. Latent Semantic Analysis

To complement our topic modelling analysis, we now examine how the similarity of central bank research to frontier research in macroeconomics has evolved over time. Our prior is that if central bank researchers have indeed shifted away from a narrow focus on macro modelling over time, then we should expect the 'distance' of central bank research to frontier research dedicated to macroeconomic modelling to have increased. The methodology used here closely follows Iaria et al. (2018), who used LSA to examine the similarity of academic research between opposing camps involved in World War 1.

LSA is a machine learning technique that uncovers semantic connections between words. It was developed by Deerwester et al. (1990) for the task of automatically retrieving information from search queries. LSA improved search results by taking into account the relationships and potential multiple meanings of words. In our context, using this algorithm means that research abstracts with completely different words can still be classified as similar if the words are regularly used in similar contexts. A neat example is provided by Landauer (2007). Using LSA, the passages *a circle's diameter* and the *radius of spheres* have similar meaning, despite having no word in common. By contrast, the text *music of the spheres* is measured as dissimilar by LSA. This property makes LSA much more preferable than using raw word frequencies to make comparisons.

LSA learns the hidden semantic connection between words and documents by using Truncated Singular Value Decomposition (TSVD) – the same transformation used in Principal Component Analysis. TSVD reduces the dimensionality of our preprocessed document-term matrix $(D \times V)$ to a document-component matrix $(D \times C)$. The number of components is user-chosen. Including both central bank research abstracts and academic abstracts, our document-term matrix has dimension of around $8500 \times 14000$, which reduces to a document component matrix of $8500 \times 300$.[3] The rows of this latter matrix capture the semantic relationships between documents (for a good overview see Iaria et al. (2018)'s technical Appendix as well as Schwarz (2017)).

With the document component matrix at our disposal, we use it to measure the 'similarity' of central bank research to frontier macroeconomic research in the following steps:

1. First, we build a 'look-up index' using only the research from the academic literature – namely, *The Journal of Macroeconomics* and *The B.E. Journal of Macroeconomics*.

[3]Our results are robust to various component sizes. Following Iaria et al. (2018) we settled on 300 components.

2. Second, for every central bank research abstract we calculate two similarity measures to this index:
   (a) the similarity to the most similar abstract from the academic literature; and
   (b) the average similarity for the five most similar abstracts.

The similarity of abstracts is calculated using the cosine similarity – often used in machine learning. The cosine similarity of central bank research abstract $i$ to an academic abstract $j$ is given by the following expression:

$$\frac{\sum_{c=1}^{C} \delta_{i,c}\delta_{j,c}}{\sqrt{\sum_c \delta_{i,c}^2}\sqrt{\Sigma_c \delta_{j,c}^2}},$$

where $\delta_{i,c}$ and $\delta_{j,c}$ are the elements inside the document-component matrix for documents $i$ and $j$.

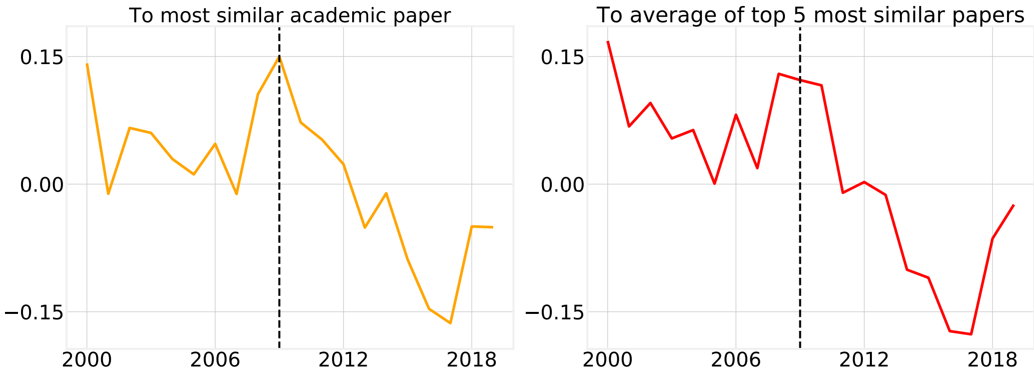
The similarity score is 1 for titles that are identical and 0 for titles that are completely different. We standardize the similarity measures to have zero mean and unit variance. An example of a randomly chosen central bank research abstract and its most 'similar' academic counterpart is provided in Appendix E.

### *4.1. LSA results*

By selecting the best academic match for every central bank abstract (as defined by the cosine similarity) we can examine trends in abstract similarities over time.

After the crisis, the similarity between central bank research and that published in the top macroeconomic journals declined by around 0.2 standard deviations (left panel, Figure 6). Moreover, these results are unchanged if we measure abstract similarity using the five most similar abstracts from the academic literature (right panel, Figure 6).

**Figure 6: Standardized Cosine Similarity to Frontier Research**




Notes: Similarities are standardized to have zero mean and unit variance
Sources: authors' calculations; IDEAS

To examine if these shifts are significant and which central banks drove the change we estimate Equation 4 below:

$$S_{ij,t} = \sum_{t=1}^{T=2} \sum_{c=1}^{C=9} \lambda_{c,t} D_{c,t} + \varepsilon_{c,t}. \tag{4}$$

Here $S_{ij,t}$ denotes the similarity of central bank research abstract $i$ to its most similar academic abstract $j$ in time period $t$ (pre and post crisis). $D_{c,t}$ is a dummy variable for central bank $c$ in time period $t$. With this simple regression, we can examine if the cosine similarity shifted significantly in the post-crisis period for each central bank.

The results show that the divergence from frontier research has been driven by the US Federal Reserve, the ECB and Sweden's Riksbank. In the pre-crisis period, the Fed and the ECB were significantly more similar than other central banks to frontier macroeconomic research, with cosine similarities around 0.1 standard deviations above the mean (Table 6, top panel). However, post crisis, similarities declined significantly, by around 0.13–0.14 standard deviations. The decline for the Riksbank was even more pronounced, falling by around 0.3 standard deviations (Table 6, bottom panel).

**Table 6: Pre- and Post-Crisis Research Similarities**

| | Cosine similarity |
|---|---|
| *Pre-crisis standardized cosine similarity* | |
| Fed | 0.11*** |
| ECB | 0.09*** |
| BoE | -0.08 |
| BoC | -0.03 |
| BoJ | -0.05 |
| RBA | 0.21** |
| RBNZ | 0.04 |
| Norges | 0.07 |
| Riksbank | 0.12 |
| *Change in standardized cosine similarity* | |
| post crisis × Fed | -0.13** |
| post crisis × ECB | -0.14*** |
| post crisis × BoE | -0.14** |
| post crisis × BoC | 0.01 |
| post crisis × BoJ | 0.06 |
| post crisis × RBA | -0.04 |
| post crisis × RBNZ | 0.22 |
| post crisis × Norges | -0.19 |
| post crisis × Riksbank | -0.30*** |
| Observations | 6637 |

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ estimated using HAC standard errors

Sources: authors' calculations'; IDEAS

### *4.2. Controlling for the counterfactual: difference-in-difference estimation*

These results notwithstanding, an important limitation is that we do not have a counterfactual. That is, how would central bank research have evolved relative to academic research in the absence of the financial crisis?

To control for this, we estimate a difference-in-difference regression (DiD). The 'control' group is the similarity of central bank research to other central bank research. The 'treatment group' is the similarity of central bank research to frontier macro research – where we assume 'treatment' is due to the crisis causing central bankers to look beyond their macro models. We are then interested in *the difference in the change* between our treatment and control group (i.e. the DiD).

The key assumption here is 'parallel trends'. That is, the average change in the control group (bank-to-bank research similarities) should represent the counterfactual change in the treatment group (bank-to-academic similarities) if the financial crisis did not occur. In Appendix F we check the parallel trends assumption by testing for differences in the pre-crisis trends of the control and treatment research similarities. The results lend confidence that our counterfactual is valid.

The difference-in-difference regression is estimated in the following regression:

$$S_{ik,t} = \beta_0 + \beta_D D_i + \beta_{\text{Post}} \text{Post}_t + \beta_{DD} \left( D_i \times \text{Post}_t \right) + u_{it}, \tag{5}$$

where $S_{ik,t}$ is the similarity of central bank research abstract $i$ to all other abstracts $k$, where $i \neq k$. $D_i$ is a dummy variable equal to 1 for treated observations (i.e. central bank-to-academic research similarities); $Post_t$ denotes a dummy equal to 1 in the post crisis period ($year > 2008$) and 0 otherwise. The four parameters can be interpreted as follows:

| | |
|---|---|
| $\beta_0$ | :mean abstract similarity in control group before the crisis |
| $\beta_{\text{Post}}$ | :post-crisis change in the mean similarity in the control group |
| $\beta_D$ | :difference in mean similarity between treatment and controls before the crisis |
| $\beta_{DD}$ | :difference in the mean change between treatment and controls (the DiD) |

The results are presented in Table 7. The key DiD estimate ($\beta_{DD}$) shows that after the crisis, the similarity of central bank research to frontier research fell by more than the similarity between central bank research and other central bank research. That is, after differencing out other factors that may have affected the similarity of research abstracts, the crisis appears to have led to a decline in the similarity between central bank and frontier macro research. This result is consistent with central bank learning; in particular, placing less emphasis on macro modelling in the aftermath of the crisis.

Turning to the other estimates, unsurprisingly, the estimate on $\beta_D$ shows that before the crisis, central bank research abstracts were more similar to each other relative to frontier macro research (by 1 standard deviation). And the estimate on $\beta_{\text{Post}}$ shows that the similarity between central bank research declined in the post-crisis period, but the fall was not significant at standard levels (p-value$= 0.101$).

**Table 7: Post-Crisis Similarities to Frontier Macro Research**
Difference-in-difference estimates

| | Coefficient estimate | Coefficient |
|---|---|---|
| $D_i \times \text{Post}_t$ | **–0.06**$^{*}$ | $\beta_{DD}$ |
| Post-crisis | –0.037 | $\beta_{\text{Post}}$ |
| $D_i$ | –0.91$^{***}$ | $\beta_D$ |
| Constant | 0.49$^{***}$ | $\beta_0$ |
| R-squared | 0.22 | |
| Observations | 13,274 | |
| Observations | 6637 | |

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$ estimated using HAC standard errors;
the number of observations reflect two similarities per central bank abstract: (1) the highest similarity to other central bank abstracts; and (2) the highest similarity to frontier research
Sources: authors' calculations'; IDEAS

## 5. Trying to isolate the impact of the crisis

The methods employed so far have not been able to adequately distinguish between longer-term changes in the intellectual debate versus something directly attributable to the financial crisis.

In this section, we attempt to make this distinction in a regression setting similar to regression discontinuity design (see Angrist and Pischke (2008), Ch. 6). In particular, we try to isolate the effect of the crisis on the topic proportions estimated for each research abstract in Section 3.

The idea behind this approach is that research published just before the onset of the crisis and shortly after the crisis should be very similar, except for the impact of the crisis. This can then be exploited for identification. We define our 'control group' as research published in 2007 and the 'treatment group' as research published in 2011. We use a three-year 'discontinuity window' to capture lags in the learning process as well as publication delays.[4]

Of course, an important assumption is that other factors that affect research trends – for example, senior staff in the central bank's research department that set research agendas

[4]The results presented here are broadly similar if we use a two-year window (2008-09) or a one-year window (2008).

as well as the ability to access big data sets – should show no jumps within the three-year window.

Specifically, we want to estimate the impact of the crisis in the following equation:

$$T_i = \beta + \alpha_i D_i + u_i, \tag{6}$$

where $T_i$ denotes the topic proportion for topic $i$ (here we omit the abstract subscript for simplicity). Treatment status $D_i$ depends on a whether the research was published before or after the crisis. That is:

$$D_i = 1\left(T_i > \overline{year}\right) \text{ (or viceversa).}$$

The impact of the crisis (the so-called treatment effect) is then identified by the following expression, which captures the expected value of the topic proportion just before and after the crisis.

$$\lim_{year \downarrow \overline{year}} E(T|year) - \lim_{year \uparrow \overline{year}} E(T|year).$$

This is quite intuitive: if the only reason why research in 2007 was different to research in 2011 was the effect of the crisis, then the impact of treatment will be the difference in outcomes between observations in 2007 compared to 2011.

We can estimate $\alpha_i$ in Equation 6 using a flexible control function approach:

$$E\left(T_i|D_i, year_i\right) = \beta + E\left(\alpha_i|\overline{year}\right) D_i + k\left(year_i\right),$$

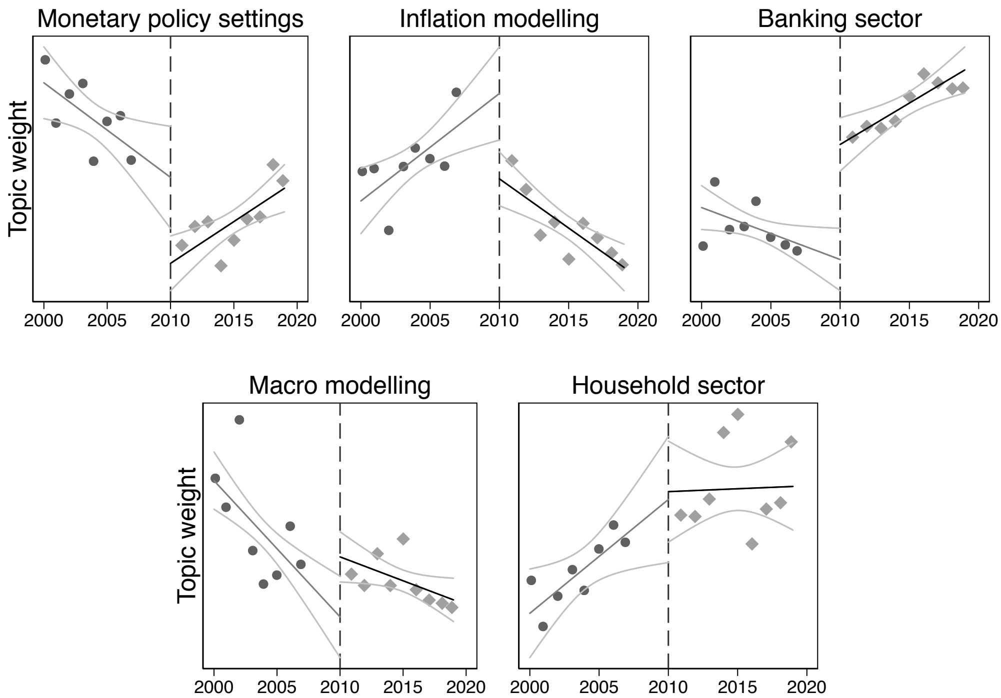
where $k\left(year_i\right)$ is the control function, which can be approximated using polynomials, such as:

$$T_i = \beta + E\left(\alpha_i|\overline{year}\right) D_i + \gamma_1 year_1 + \gamma_2 year_i^2 + \ldots + \gamma_p year_i^p + \epsilon_i \tag{7}$$

Figure 7 plots the discontinuities we identify for the five topics whose weights exhibited meaningful variation over time. Table 8 presents the estimates of $\alpha_i$ from Equation 7 using a $4^{th}$-degree polynomial.

The results indicate that the crisis meaningfully reduced the amount of research dedicated to the topics *monetary policy transmission* and *inflation modelling*. Understandably, the crisis also meaningfully increased the amount of research dedicated to the *banking sector*.

**Figure 7: Trying to Identify the Impact of the Crisis**
A Regression Discontinuity Design




Notes: The Figure clusters topic proportions by year (with each dot representing all the research produced in that year) and fits a linear trend for the pre-crisis (before 2008) and post-crisis period (after 2010); the years 2008–10 are excluded to capture lags in the learning process as well as publication delays; the discontinuity used for identification is indicated by the vertical line; the bands indicate a 95 % confidence interval

Sources: authors' calculations; IDEAS

Interestingly, the results from this exercise indicate that the amount of research dedicated to the *household sector* and *macro modelling* was not significantly affected by the crisis. This is despite the significant trends in their topic proportions over the entire 20-year sample period identified in Section 3.

Two explanations can help square these findings. First, there was a significantly increasing pre-crisis trend in the *household sector* topic. This speaks to a long-run increase in the ability to create, access and analyse micro-level datasets over our 20-year sample period (i.e. the so-called 'explosion' of big data). This in turn is likely to have contributed to a trend increase in the topic proportion. Likewise, there was a significantly decreasing pre-crisis trend in the *macro modelling* topic, perhaps reflecting a disenfranchisement with the usefulness of large-scale macro models even before the onset of the financial crisis.

Second, lessons from the crisis may be long-lived. Our regression discontinuity design only captures more immediate changes related to the crisis. For example, it is of little surprise that the regression discontinuity design identifies the banking sector as being the topic most affected by the financial crisis. Leverage in the financial system was at the heart of the financial crisis and highly-leveraged financial institutions (e.g. Bear Stearns and Lehmans) played a fundamental role in significantly amplifying the crisis. Learning from this experience was obviously front-of-mind for monetary-policy researchers following the crisis.

**Table 8: A Regression Discontinuity Design**
Trying to Identify the Impact of the Crisis

| | **Proportions (%)** | | **Change (ppt)** |
|---|---|---|---|
| *Topic* | *2007* | *2011* | *Effect* |
| Monetary policy transmission | 0.13 | 0.08 | -0.04** |
| Inflation modelling | 0.19 | 0.15 | -0.04** |
| Macro modelling | 0.11 | 0.11 | -0.01 |
| Banking sector | 0.08 | 0.16 | 0.08*** |
| Household sector | 0.05 | 0.05 | 0.01 |

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sources: authors' calculations; IDEAS

## 6. Conclusion

In this paper, we scraped the IDEAS economic research repository to gather around 6600 research abstracts written by advanced economy central banks over the past 20 years. With these data, we examined how the intellectual debate occurring within central banks has evolved, with a focus on the impact of the financial crisis.

In the context of the financial crisis, we focused on examining two so-called 'lessons learned': (1) that central banks failed to pick up on early-warning indicators prior to the crisis because they relied too heavily on the forecasts from macro models; and (2) that the usefulness of micro data in teasing out the causes of aggregate fluctuations was under-appreciated.

Our main conclusion is that these identified shortcomings appear to have contributed to subsequent changes in the intellectual debate occurring within central banks.

Using a topic model, we identified significant shifts in a number of fixed research topics over time – as well as changes within topics over time. Notably, there was an increase in research dedicated to better understanding the behaviour of individual sectors of the economy, namely the banking and household sectors. We also identified a decline in research that focused on modelling (for example, macro modelling and inflation modelling.) This notwithstanding, some of these trends were already occurring well before the financial crisis.

Using LSA, combined with a difference-in-difference estimation, we also identified a post-crisis fall in the similarity between central bank research and academic research dedicated to macro modelling.

To try and isolate the direct impact of the crisis on these trends (as compared to longer-term structural drivers) we also performed a type of regression discontinuity design. After controlling for existing trends in the data, there was a meaningful fall in the amount of intellectual effort dedicated to modelling inflation and an increase in research on the banking sector.

However, these results should not be over interpreted and are subject to a number of limitations. First, in our simple LDA model (used to examine the change in fixed topics over time), we assumed that documents were exchangeable throughout the 20-year period (that is, the distribution over words that defines a topic was assumed to be fixed). This is (arguably) an unreasonable assumption given our sample period covers two decades of research. Furthermore, future work would benefit from making better use of the financial crisis as a natural experiment to disentangle the impact of long-run versus short-run drivers of changes in research trends (although we took several steps in this direction).

Future research could also make more use of other meta-data associated with the research abstracts we collected. For example, author-topic models could be used to identify 'superstar' researchers that were ahead of the learning curve and drove changes in the intellectual debate.

Finally, the next step in a broader research agenda would be to map the changes we identified into changes in the policy debate, by examining more policy-relevant text data, such as speeches, minutes and quarterly monetary policy statements. This would speak to the 'influence' economic research is having in the policy sphere, which first and foremost is the goal of any central banking research department.

## Appendix A. LDA

### *Appendix A.1. LDA as a probabilistic model*

The LDA topic model is most easily described by the imaginary random process by which the model assumes documents arise from:

1. First, authors begin by choosing a distribution over topics for their paper.
2. Then, before writing a word, authors choose what topic it belongs to.
3. Finally, authors look up the distribution over terms associated with that topic, draw a word from the topic and commit that word to paper.

To generate a complete document, this process is repeated for every word. A new document is then created, with words chosen in the same way. Importantly, the topics stay the same from document to document, but how each document exhibits those topics changes.

This is the hallmark of the LDA model, which is formally a mixed membership model. That is, each document is coming from a mixture model, where the mixture proportions change from document-to-document, but the mixture components are fixed across the whole collection.

The machine learning and algorithmic problem is that we do not get to observe the above imaginary generative process. Instead, to infer all of the values associated with this latent generative process, posterior inference is used – that is, the *conditional distribution* of the hidden variables given the observed variables. The observed variables are the words of the documents. The hidden variables are the per-document topic proportions; the per-topic word distributions; and the per-word topic assignment.

We can describe LDA more formally with the following notation (for neat explanation, see Blei (2012)). The topics are $\beta_{1:K}$, where each $\beta_k$ is a distribution over words. The topic proportions for the $d^{th}$ document are $\theta_d$ where $\theta_{d,k}$ is the topic proportion for topic $k$ in document $d$. The topic assignments for the $d^{th}$ document are $z_d$, where $z_{d,n}$ is the topic assignment for the $n^{th}$ word in document $d$. Finally, the observed words for document $d$ are $w_d$, where $w_{d,n}$ is the $n^{th}$ word in document $d$, which is an element from the fixed vocabulary.

The generative process for LDA then corresponds to the following joint distribution of the observed and hidden variables:

$$
\begin{aligned}
&p\left(\beta_{1:K}, \theta_{1:D}, z_{1:D}, w_{1:D}\right) \\
&= \prod_{i=1}^{K} p\left(\beta_i\right) \prod_{d=1}^{D} p\left(\theta_d\right) \left(\prod_{n=1}^{N} p\left(z_{d,n}|\theta_d\right) p\left(w_{d,n}|\beta_{1:\kappa}, z_{d,n}\right)\right)
\end{aligned}
$$

This distribution specifies a number of dependencies. For example, the topic assignment $z_{d,n}$ depends on the per-document topic proportions $\theta_d$. As another example, the observed word $w_{d,n}$ depends on the topic assignment $z_{d,n}$ and all of the topics $\beta_{1:K}$.

*Appendix A.2. Priors and posterior computation for LDA*

In a Bayesian inference context, the posterior distribution gives the probability of observing certain parameter values after observing the data (the corpus in the LDA case). This is in contrast to the prior distribution, which assigns probabilities before considering the data.

Relating this to LDA, the prior distributions are usually selected to be Dirichlet distributions due to the fact that these distributions lend high probability to a small group of words. The supports the underlying intuition that topics are characterised by frequent occurrences of a small number of words.

Additionally, a Dirichlet also allows simpler calculation of the posterior due to its conjugate relationship with the multinomial distribution. This implies that the prior and the posterior are from the same distributional family and we can write down a closed form expression for the posterior.

These two factors have ensured that a Dirichlet prior has been the standard approach in LDA. Furthermore, symmetrical distributions are usually used to signal that all topics have equal chance of being assigned to a document and all words have equal chance of being assigned to a topic. This results in the most general and flexible formulation of an LDA approach.

Formally stated, we take:

$$
\theta_d \sim \mathrm{Dir}(\alpha) \qquad\qquad \beta_k \sim \mathrm{Dir}(\phi)
$$

where $\alpha$ and $\phi$ are the hyper-parameters for our priors.

As mentioned above, we cannot actually observe the imaginary generative process which has been specified and the hidden parameter values must be inferred from posterior analysis. Algebraically, the posterior can be denoted as follows:

$$p(\theta, \mathbf{z}|\mathbf{w}, \alpha, \beta) = \frac{p(\theta, \mathbf{z}, \mathbf{w}|\alpha, \beta)}{p(\mathbf{w}|\alpha, \beta)}$$

That is, we would like to find the distribution of the hidden latent variables $\theta$ and $z$, given our prior distributions and the observed data.

Unfortunately this expression is known to be intractable to compute and must be approximated. One approach is to use *variational inference* where Jensen's inequality is used to repeatedly find a lower bound for the log-likelihood (this approach is used here). Another is to use Gibbs sampling, which constructs a sequence of random variables dependent on the previous (i.e. a *Markov chain*), whose limiting distribution is the posterior.

### *Appendix A.3. Dynamic topic models*

Dynamic topic models extend the idea of LDA and introduce a time dependency element.

Documents are grouped into time periods and the given topics in each sequential time period are related. That is to say that the topics cannot wildly vary from one year to the next (assuming we were grouping by year).

It is important to note that the overall number of topics in the model remains fixed. To add dynamics to the model, we introduce a probabilistic relationship between the topic distribution in one period and the topic distribution in the other period.

Formally stated:

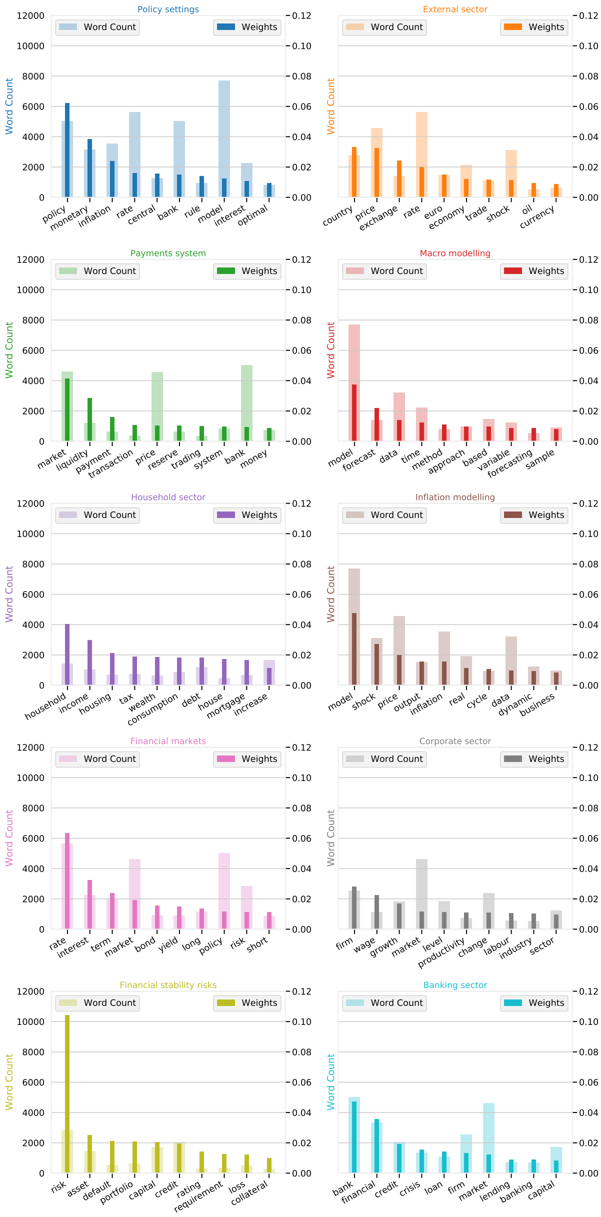
$$\beta_{t,k}|\beta_{t-1,k} \sim N\left(\beta_{t-1,k}, \sigma^2 I\right)$$
$$\alpha_t|\alpha_{t-1} \sim N\left(\alpha_{t-1}, \delta^2 I\right)$$

where $\beta_{t,k}$ is the word distribution of topic $k$ at time $t$ and $\alpha_t$ is the per-document topic distribution at time $t$.

Aside from the added dynamics, the LDA formulation and computation operates very similarly within each time period. However, an important difference is that the Dirichlet is not amenable to sequential modelling. Instead, the natural parameters of each topic $\beta_{t,k}$ are chained together in a state space model that evolves with Gaussian noise.

# Appendix B. Importance of Topic Keywords

## Figure B.8: Word Count and Importance of Topic Keywords



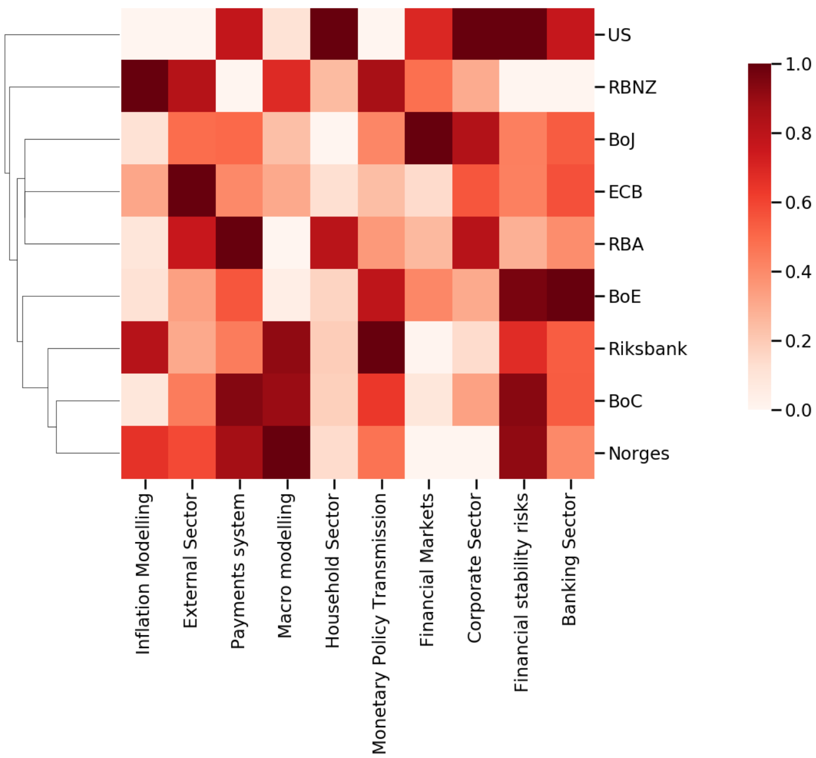

Sources: authors' calculations; IDEAS

## Appendix C. Relative topic weights between central banks

Figure F.10 shows a heatmap of the relative topic strength for each central bank. It also shows a dendrogram obtained from hierarchical clustering indicating which central banks are most similar in terms of the relative weight they attach to different research topics.

As noted in the text, the US Federal Reserve is somewhat of an outlier. It places relatively less weight on topics dedicated to *inflation modelling*; *macro modelling*; and *monetary policy transmission*. Instead, the Fed has traditionally placed more emphasis on individual sectors of the economy as well as financial stability risks. The hierarchical clustering shows that the Fed is least similar to other central banks in terms of the topics it focuses on.

**Figure C.9: The Relative Importance of Each Topic**
By Central Bank




Sources: authors' calculations; IDEAS

## Appendix D. Dynamic Topics

**Table D.9: The Evolution of the Household Sector Topic**

| **2000** | | **2009** | | **2018** | |
|---|---|---|---|---|---|
| **words** | **weights** | **words** | **weights** | **words** | **weights** |
| rate | 0.071 | household | 0.036 | household | 0.047 |
| exchange | 0.054 | rate | 0.033 | housing | 0.029 |
| income | 0.028 | exchange | 0.030 | income | 0.028 |
| household | 0.026 | income | 0.026 | rate | 0.025 |
| consumption | 0.024 | tax | 0.023 | wealth | 0.020 |
| tax | 0.020 | consumption | 0.020 | mortgage | 0.019 |
| currency | 0.019 | housing | 0.018 | exchange | 0.018 |
| wealth | 0.018 | wealth | 0.017 | tax | 0.017 |
| housing | 0.012 | currency | 0.014 | consumption | 0.017 |
| saving | 0.011 | house | 0.014 | house | 0.012 |
| dollar | 0.010 | mortgage | 0.013 | data | 0.012 |
| mortgage | 0.009 | individual | 0.013 | debt | 0.010 |
| individual | 0.008 | data | 0.009 | individual | 0.010 |
| house | 0.008 | estimate | 0.009 | currency | 0.010 |
| data | 0.008 | debt | 0.008 | level | 0.009 |
| estimate | 0.007 | import | 0.007 | share | 0.009 |
| import | 0.007 | level | 0.007 | estimate | 0.009 |
| share | 0.007 | dollar | 0.007 | price | 0.007 |
| level | 0.006 | price | 0.007 | saving | 0.006 |
| pas | 0.006 | share | 0.007 | home | 0.006 |
| price | 0.006 | percent | 0.006 | survey | 0.006 |
| debt | 0.005 | saving | 0.006 | inequality | 0.006 |
| panel | 0.005 | home | 0.006 | percent | 0.006 |
| age | 0.005 | elasticity | 0.006 | dollar | 0.005 |
| account | 0.005 | effect | 0.005 | effect | 0.005 |
| elasticity | 0.005 | spending | 0.005 | distribution | 0.005 |
| percent | 0.004 | account | 0.005 | elasticity | 0.005 |
| ratio | 0.004 | value | 0.005 | import | 0.005 |
| home | 0.004 | distribution | 0.005 | value | 0.005 |
| effect | 0.004 | emerging | 0.005 | across | 0.005 |

Table D.10: The Evolution of the Macro Modelling Topic

| 2000 | | 2009 | | 2018 | |
|---|---|---|---|---|---|
| **words** | **weights** | **words** | **weights** | **words** | **weights** |
| model | 0.051 | model | 0.067 | model | 0.047 |
| output | 0.014 | time | 0.009 | time | 0.010 |
| inflation | 0.011 | data | 0.009 | estimate | 0.009 |
| estimate | 0.010 | estimate | 0.009 | approach | 0.008 |
| gap | 0.009 | dynamic | 0.008 | equilibrium | 0.008 |
| equilibrium | 0.009 | equilibrium | 0.008 | method | 0.008 |
| data | 0.008 | approach | 0.007 | data | 0.008 |
| approach | 0.008 | new | 0.007 | based | 0.007 |
| variable | 0.007 | method | 0.007 | dynamic | 0.006 |
| test | 0.007 | based | 0.006 | new | 0.006 |
| method | 0.007 | variable | 0.006 | sample | 0.006 |
| optimal | 0.007 | inflation | 0.006 | variable | 0.006 |
| estimated | 0.007 | show | 0.006 | parameter | 0.005 |
| dynamic | 0.007 | estimated | 0.006 | used | 0.005 |
| new | 0.006 | parameter | 0.006 | show | 0.005 |
| based | 0.006 | sample | 0.005 | distribution | 0.005 |
| time | 0.006 | used | 0.005 | gap | 0.005 |
| used | 0.005 | structural | 0.005 | framework | 0.004 |
| curve | 0.005 | gap | 0.004 | estimation | 0.004 |
| parameter | 0.005 | dsge | 0.004 | test | 0.004 |
| sample | 0.005 | estimation | 0.004 | estimated | 0.004 |
| empirical | 0.005 | output | 0.004 | function | 0.004 |
| show | 0.005 | test | 0.004 | state | 0.004 |
| function | 0.004 | optimal | 0.004 | inflation | 0.004 |
| estimation | 0.004 | well | 0.004 | empirical | 0.004 |
| targeting | 0.004 | empirical | 0.004 | well | 0.004 |
| uncertainty | 0.004 | general | 0.004 | provide | 0.004 |
| structural | 0.004 | function | 0.004 | different | 0.004 |
| different | 0.004 | var | 0.004 | non | 0.004 |
| well | 0.004 | keynesian | 0.004 | dsge | 0.003 |

## Appendix E. Document similarity example

**Central Bank**: **Skaperdas (2017)**, 'How Effective is Monetary Policy at the Zero Lower Bound? Identification Through Industry Heterogeneity', *Federal Reserve System*

US monetary policy was constrained from 2008 to 2015 by the zero lower bound, during which the Federal Reserve would likely have lowered the federal funds rate further if it were able to. This paper uses industry-level data to examine how growth was affected. Despite the zero bound constraint, industries historically more sensitive to interest rates, such as construction, performed relatively well in comparison to industries not typically affected by monetary policy. Further evidence suggests that unconventional policy lowered the effective stance of policy below zero.

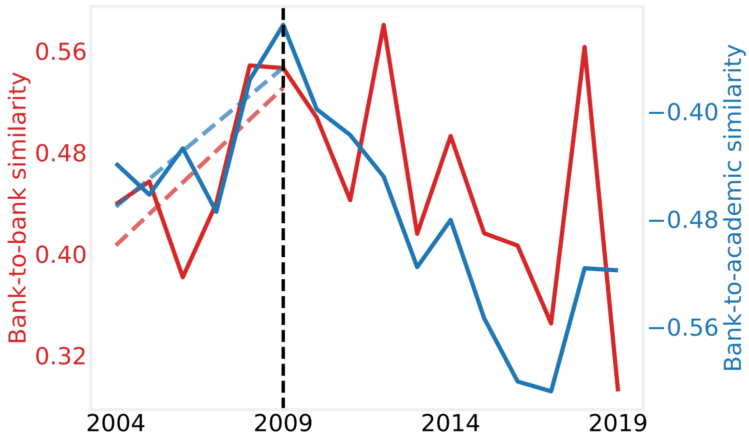
**Academia: Belongia, M and Ireland P (2017)**, 'Circumventing the zero lower bound with monetary policy rules based on money', *Journal of Macroeconomics*

Discussions of monetary policy rules after the 2007–2009 recession highlight the potential ineffectiveness of a central bank's actions when the short-term interest rate under its control is limited by the zero lower bound. This perspective assumes, in a manner consistent with the canonical New Keynesian model, that the quantity of money has no role to play in transmitting a central bank's actions to economic activity. This paper examines the validity of this claim and investigates the properties of alternative monetary policy rules based on control of the monetary base or a monetary aggregate in lieu of the capacity to manipulate a short-term interest rate. The results indicate that rules of this type have the potential to guide monetary policy decisions toward the achievement of a long-run nominal goal without being constrained by the zero lower bound on a nominal interest rate. They suggest, in particular, that by exerting its influence over the monetary base or a broader aggregate, the Federal Reserve could more effectively stabilize nominal income around a long-run target path, even in a low or zero interest-rate environment.

## Appendix F. The parallel trends assumption

In this Appendix we check the parallel trends assumption underlying the DiD results presented in Table 7. The parallel trends observed in the pre-treatment paths of the treatment and control groups lends some confidence that the control group provides a good counterfactual. However, we cannot directly test this identifying assumption, which is, by construction, untestable.

**Figure F.10: The Parallel Trends Assumption**




Notes: The left-hand axis shows the standardised cosine similarity (i.e. mean 0 and standard deviation 1) between central bank research (i.e. the control group); the right-hand axis shows the standardised cosine similarity between central bank research and frontier macro research (i.e. the treatment group)

Sources: authors' calculations; IDEAS

## References


Angrist, J. and J.-S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Beyer, A., B. Coeure, and C. Mendicino (2017). Foreword – The crisis, Ten Years After: Lessons Learnt for Monetary and Financial Research. *Economie et Statistique / Economics and Statistics* (494-495), 45–64.

Blei, D. and J. Lafferty (2006). Dynamic Topic Models. In Proceedings of the 23rd International Conference on Machine Learning.

Blei, D., A. Ng, and M. Jordan (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research 3*, 993–1022.

Blei, D. M. (2012). Probabilistic Topic Models. *Commun. ACM 55*(4), 77–84.

Braude, J., Z. Eckstein, S. Fischer, and K. Flug (2013). *The Great Recession: Lessons for Central Bankers*. Mit Press.

Chang, J., S. Gerrish, C. Wang, J. Boyd-Graber, and D. Blei (2009). Reading Tea Leaves: How Humans Interpret Topic Models. In Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22*, pp. 288–296. Curran Associates, Inc.

Deerwester, S., S. Dumais, G. Furnas, T. Landauer, and R. Harshman (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science 41*(6), 391–407.

Fligstein, N., J. Brundage, and M. Schultz (2014). Why the Federal Reserve Failed to See the Financial Crisis of 2008: The Role of 'Macroeconomics' as a Sense making and Cultural Frame. Institute for Research on Labor and Employment, Working Paper Series, Institute of Industrial Relations, UC Berkeley.

Friedman, M. and A. Schwartz (1963). *A Monetary History of the United States, 1867–1960*. National Bureau of Economic Research, Inc.

Hansen, S., M. McMahon, and A. Prat (2017). Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach. *The Quarterly Journal of Economics 133*(2), 801–870.

Iaria, A., C. Schwarz, and F. Waldinger (2018). Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science. *The Quarterly Journal of Economics 133*(2), 927–991.

Kohn, D. and B. Sack (2018). Monetary Policy during the Financial Crisis. Presented at the Brookings and Yale School of Management event on Financial Stability, *Responding to the Global Financial Crisis: What We Did and Why We Did It*, Washington, D.C.

Landauer, T. (2007). LSA as a Theory of Meaning. *Handbook of Latent Semantic Analysis*, 3–34.

Lee, H. and P. Kang (2018). Identifying Core Topics in Technology and Innovation Management Studies: A Topic Model Approach. *The Journal of Technology Transfer 43*(5), 1291–1317.

Mian, A. and A. Sufi (2010). The Great Recession: Lessons from Microeconomic Data. *American Economic Review 100*(2), 51–56.

Potter, S. (2019). Models Only Get You so Far. Remarks for the Federal Reserve Banks of Atlanta and New York's First Annual Joint Research Day.

Schwarz, C. (2017). lsemantica : A Stata Command for Text Similarity based on Latent Semantic Analysis. *The Stata Journal*.

Simons, J. (2019). Ten Years of Research – What Have We Learnt Since the Financial Crisis? Remarks to the Economic Society of Australia (QLD) and Griffith University Symposium.

Tetlock, P. E. and D. Gardner (2015). *Superforecasting: The Art and Science of Prediction*. New York, NY, USA: Crown Publishing Group.

van der Maaten, L. and G. Hinton (2008). Visualizing Data using t-SNE . *Journal of Machine Learning Research 9*, 2579–2605.