Changing inflation perception and the propagation of news shocks

Torsten Schmidt, Henrik Müller, Jonas Rieger, Tobias Schmidt, Carsten Jentsch*

*Torsten Schmidt is Head of the Macroeconomics and Public Finance department at RWI and professor for Empirical Macroeconomics at Ruhr-University Bochum. Henrik Müller holds the chair of Economic Policy Journalism at TU Dortmund University and is a co-founder of Dortmund Center for data-based Media Analysis (DoCMA). Tobias Schmidt and Jonas Rieger are researchers at TU Dortmund University and DoCMA. Carsten Jentsch is a professor for business and social statistics at TU Dortmund University and member of DoCMA. The authors would like to thank Almut Balleer, Boris Blagov, Lena Marie Hufnagel and Nico Hornig for helpful comments and ideas.

Abstract

In this paper, we analyze the information content of inflation-related media coverage. Drawing on a previously established indicator, the Inflation Perception Indicator (IPI), we examine the salience and framing of inflation in a corpus of 50,000 articles published by major German newspapers between 2001 and 2024. The IPI captures the share of inflation-related reporting over time and is decomposed into dynamic thematic sub-indicators using the RollingLDA algorithm, a time-consistent dynamic topic modeling approach. To analyze the information content of these new indicators, we estimate a threshold BVAR for variables from a New Keynesian macroeconomic model. Our results lead to two major insights: 1) It is not just the gross intensity of coverage that shapes expectations, but more so the subtler facets of reporting. The effects on inflation perception critically hinge on the specific issues and frames highlighted by the media. Our results show remarkably different response patterns for different kinds of inflation shocks, which our methodology ably captures. In particular, we find that shocks related to Raw Material prices show effects like a demand shock, while news shocks spilling over from abroad resemble the pattern of supply shocks. 2) Nonetheless, the sheer volume of public attention matters. In an environment of intensive inflation reporting, the effects on expectations are decisively stronger than in a low inflation regime. Both insights have potentially significant implications for the conduct of monetary policy.

JEL: C22, C82, D84, E31

Key words: Attention Cycle, Expectations, Inflation, Media, Perception

1. Introduction

In many situations, economic agents incorporate information on current and future price trends into their decisions. Several channels are available for this purpose, but these are associated with different costs for the various actors. While companies obtain effective price information by directly observing their sales and procurement markets, private households are likely to be more dependent on media reports. However, the media do not report consistently on certain topics, but rather on specific events. These different intensities of media reporting are accompanied with different levels of attention by private households. Recent papers by Flynn and Sastry (2024) and Pfäuti (2024) find that the level of attention has a substantial effect on the strength of the transmission of macroeconomic shocks.

In this paper, we highlight the event-driven nature of news: at any point in time, the public sphere tends to be dominated by a limited number of issues, that follow distinct attention cycles (Downs, 1972; Lörcher and Neverla, 2015; Waldherr, 2014). In modern-day media environments, these patterns can amount to rather extreme outbursts of public attention encompassing not only traditional media, but also social media and other parts of the public sphere, a phenomenon Waisbord and Russell (2020) refer to as "news flashpoints". The media stress what's new and problematic, but attention wanes as newness fades, no matter if the underlying problem passes or persists.

This view on the role of media can be linked to the concept of rational inattention by using the model of Nimark and Pitschner (2019). These authors argue that businesses and households delegate information gathering and viewing to media companies. Media companies gather information and monitor inflation developments and decide on a case-by-case basis whether to report on current developments. In this regard the reporting of media is state dependent. By reporting on current developments, they give readers a signal that a new development potentially important to them has begun. The media not only report on the development of the inflation rate, but also on its causes and possible

consequences. Based on this signal households have the choice to collect information about price developments and adjust their inflation expectations accordingly.

To analyze the information content in media reporting in Germany, we use indicators for different newspaper topics related to inflation developments.[1] Our hypothesis is that not only the information intensity is important for the effect of news on inflation expectations, but also the sources of inflation. For this reason, in this paper, we make use of a measure called Inflation Perception Indicator (IPI) that was first introduced by Rieger et al., 2023. Based on large text corpora of newspaper articles, it is calculated by applying the dynamic topic modeling method RollingLDA (Rieger et al., 2021), which refines the original (static) LDA algorithm (Blei et al., 2003) by allowing for temporal changes in the topics when adding data. This approach enables us to sequentially construct time-consistent indicator time series.

The IPI is designed to capture elements of inflation narratives floating in the public sphere, thereby complementing established measures of inflation expectations derived from surveys and financial market data. Since the price momentum is largely driven by inflation expectations, identifying narrative factors that drive these projections strikes us as a worthwhile endeavor. Based on the RollingLDA method, nine inflation-related topics are identified. These include the coverage of central banks, fiscal and re-distributional policies, raw material prices and others. The time series for these topics show that the intensity and composition of reporting varies considerably over time.

In the subsequent empirical study, we use a threshold BVAR model to examine the information content of inflation topics for macroeconomic development. This model allows us to distinguish two different inflation environments. The approach of rational inattention states that the level of inflation influences the level of attention and therefore the formation of inflation expectations. The hypothesis is that rational inattention is overcome by the sheer volume of reporting, prompting economic agents to watch news media topics closely and adjust their expectations accordingly. As attention peters out, economic

[1] Our approach is similar to Angelico et al. (2022).

agents revert to inattention.

Our results show that the constructed inflation topics contain different information about the causes of price increases. For example, shocks in the topic Raw Materials (see Section 4 for details) have effects that are similar to a demand shock, while shocks in the topic of Emerging Markets, i.e., informational spillovers from abroad have effects that are similar to a supply shock. Moreover, in most cases, it is possible to distinguish two regimes for the level of inflation. In our estimation, the average level for the threshold inflation rate is 1.9%, which is similar to related studies (Korenok et al., 2023). We also find different effects of news shocks on the formation of private household expectations in the two inflation regimes. While, in the high inflation regime, news shocks of the topics Raw Materials and Emerging Market have effects on expectations of private households, this is not the case in low inflation regimes. This is in line with the finding that, in periods of intensive reporting about inflation, media play a role in the formation of inflation expectations (Lamla and Lein, 2015; Larsen et al., 2021) and that different inflation topics have different effects on inflation expectations (Dräger and Nghiem, 2021). The IPI methodology is able shifts in content over time in a reliable fashion.

This paper is organized as follows: Section 2 provides an overview of the relevant literature focusing on studies that point to a time-varying importance of media for inflation expectations. Section 3 provides a description of the RollingLDA method, the data base and our processing steps. Section 4 presents the (sub)indicator time series for the Inflation Perception Indicator and discusses specific topics identified by RollingLDA. In Section 5, we present the threshold VAR model and discuss the resulting impulse response analysis. Section 6 concludes.

2. Inflation perception and attention cycles: A review of the literature

In this study, we combine various recent methods for analysing inflation perception in the media and thereby supplementing some new findings.Newspaper reports on inflation can be important for overall economic development, because they reduce the cost of obtaining information about the macroeconomic outlook (Reis, 2006; Sims, 2003). Due to the cost of information collection, people only form and adjust economic expectations from time to time. In the meantime, expectations remain unaltered. In this context, information rigidity is an indicator for the frequency of the adjustment of expectations. The longer the time period between expectation adjustment is, the higher is the information rigidity. Coibion and Gorodnichenko (2015) found empirical evidence for the US that this information rigidity varies over time. In particular, they found that during recessions the information rigidity declines.

Bracha and Tang (2022) have taken the idea of time-varying carelessness and reviewed it for the formation of inflation expectations. To do so, they constructed a measure of inattention from the Michigan Consumer Survey data and concluded that in periods of low inflation rates the ratio is higher than in periods of high inflation, indicating that the extent of inattention is state dependent.

A rigorous empirical test of this idea is the estimation of attention thresholds to distinguish a low and a high attention regime. Korenok et al. (2023) estimate an attention threshold for inflation using a panel of countries. Attention is measured by the frequency of Google searches for “inflation”. The estimated threshold for a broad set of countries lies between 2-4%. The authors also use newspaper data for a selected set of countries to estimate attention thresholds for inflation. For Germany, the estimated threshold is 1.4% for the newspaper data compared to 1.9% for the Google data.

Pfäuti (2024) estimates an attention threshold for the US based on forecast errors and

finds that, in the high inflation regime, effects of supply shocks to inflation are twice as large as in the low attention regime.

The studies cited so far find empirical evidence that households change their attention with regard to inflation over time. However, form where they get the impetus to alter their behavior has been largely neglected in these studies. We therefore combine the threshold attention approaches to the fast growing literature on the formation of inflation expectations (D'Acunto et al., 2023; Weber et al., 2022) stressing the role of media. However, recent studies on the determinants of the formation of inflation expectations by households have differed in their assessment of the importance of media coverage. Weber et al. (2022) and D'Acunto et al. (2023), among others, provide overviews of studies on media effects with a focus on central bank communication. They summarize the results by showing that particularly the reporting of central bank actions are too complex for most households, so that they do not align their inflation expectations. Rather, households seem to rely on price signals from their environment in forming their expectations.

Nevertheless, there are a number of studies that find an effect of media coverage on the expectations of households. These studies can also provide starting points for explaining how media influence different levels of attention. Conrad et al. (2022) find media influences on the level of perceived past and expected future inflation. Users of traditional media, as opposed to social media, accurately perceive inflation rates of the recent past to be lower, while their inflation expectations for the proximate future are lower as well. Moreover, Dräger and Nghiem (2021) analyze the role of news on inflation perceptions and inflation expectations based on survey data. The authors show that households, who are able to recall news about monetary policy issues that includes inflation developments, have more accurate inflation expectations than households who cannot recall them. In addition, the authors find that the type of information has different effects. Their results suggest that information about monetary policy issues has an effect on inflation expectations. Information about financial market topics is more likely to influence expectations about nominal interest rates. It is therefore to be expected that households pay attention to

the causes of inflation. We supplement these studies by including news topics that cover different information sets.

One building block for explaining the various findings on the role of media in shaping household expectations is the way media works. Nimark and Pitschner (2019) argue that businesses and households delegate information gathering and viewing to media companies. Media companies gather information and monitor inflation developments, for example, and decide on a case-by-case basis whether to report on current developments. In this regard, the reporting of media is state-dependent and it provides information about the sources of the recent inflation developments. We use this argument for the use of the threshold model with inflation as the threshold variable.

In this regard our approach is related to Larsen et al. (2021) who take up the idea that households get information about future price trends through the media. The authors use an LDA approach and find six topics that are significantly related to private households' inflation expectations: education, transactions, health, internet and trading. The interpretation of the authors is that all of these topics are related to some sort of household expenditures. The authors find that the information rigidity is time-varying and that media coverage can explain the variation over time. Thus, through intensive coverage, media can help households adjust their inflation expectations.

The role of the media is also emphasised in the literature on central bank communication. Monetary policy makers have greatly intensified their public relations efforts, and the media have become an important transmission channel (Berger et al., 2011). As traditional monetary policy instruments such as short-term interest rates and asset purchases have been all but exhausted, forward guidance has emerged as a policy instrument in its own right, i.e., communication has become part of the standard monetary policy toolbox. Informing the public about its views concerning the workings of the economy (the model) and its current state (the data), as well as stating an explicit medium- to long-term policy goal (2 per cent annual CPI inflation), central banks intend to influence long-term yields

and thereby forward-looking decisions concerning the real economy (investment). Hansen et al. (2019) find that there is indeed such a long-term effect on asset-prices, particularly by shaping perceptions of long-run uncertainty. Ter Ellen et al. (2022) consider "narrative monetary policy surprises" and come to conclude that such information shocks lead to real macroeconomic effects. The authors stress the role of the media as information intermediaries. Seen this way, complexity-reducing media coverage can be seen as an efficient way to transmit monetary policy narratives, which, as Ter Ellen et al. (2020) point out, are "key ingredients in any interest rate decision and important for households". Thus, central banks' communicative efforts prompt media coverage and, thereby, influence expectations. Households do not need to know anything about macroeconomics and the transmission channels of monetary policy, as long as media coverage informs them of what to expect (and as long as central banks are credible and the media is trusted).

3. Data and Methods

Our content analysis is based on a corpus of three leading nation-wide German newspapers: Süddeutsche Zeitung (center left), Die Welt (center right) and Handelsblatt (business). The data was obtained from LexisNexis and from the publishing houses. Articles published between January 1, 2001 and March 31, 2024 were considered. Based on this data, we pursue two distinct but complementary analytical goals. First, we track general patterns of media attention to inflation over time. Second, we investigate how inflation is framed across thematically distinct reporting clusters, some of which may independently influence inflation expectations.

Thus, our content analysis is based on two separate but complementary components. First, we compute a simple article-share measure, the Inflation Perception Indicator (IPI), which was first introduced by Rieger et al. (2023) and tracks the relative salience of inflation-related content in the overall news corpus over time, as described above. Second, we apply a dynamic topic modeling procedure (RollingLDA) to this subset of inflation-

relevant articles to decompose the IPI into semantically coherent topics. This allows us to analyze not only how prominently inflation is covered in the media, but also how this coverage is framed and what specific narratives dominate public discourse in different periods. This procedure was first proposed by Müller et al. (2022).

3.1. Construction of an inflation perception indicator

To begin with, we construct a high-level indicator of how prominently inflation is covered in the news over time. This indicator — the Inflation Perception Indicator (IPI) — is based on a keyword-driven identification of inflation-related articles. It is defined as the monthly share of newspaper articles in the overall corpus that match a predefined set of inflation-related keywords. For each month t, we calculate:

$$\text{IPI}_t = \frac{\text{Number of inflation-relevant articles in month } t}{\text{Total number of articles in month } t}.$$

To refine a corpus of relevant articles, we excluded all articles that do not provide any information on inflation reporting — which applies to the majority of the texts in our corpus. Our goal is to consider all documents that actually mention inflation, while also including articles that address this topic without using the term "inflation". Furthermore, we aim at calibrating our indicator so that it responds early to changes in reporting. To achieve this, in a first step, we applied a broad search pattern to exclude articles unrelated to economy or prices. From this subset, we randomly sampled 300 articles and manually coded them as relevant or irrelevant. We then tested various search terms on this sample and evaluated each term's effectiveness using recall and precision metrics as proposed by Stryker et al. (2006), where recall measures a search strategy's ability to retrieve all relevant items, while precision quantifies the proportion of retrieved items that are relevant. In selecting our final search terms, we prioritized high recall over high precision, given our goal of creating a sensitive indicator that captures early changes in inflation reporting. Finally, we applied the best-performing search term in this regard to

the whole corpus.

The final search term in our analysis is of the form:

inflation OR teuerung OR geldentwertung* OR preissteigerung**[2]

It comes with a recall value of 0.808 and a precision value of 0.576. The selection of its individual components is based on both intuition and statistics. The inclusion of the word "inflation" is (intuitively argued) self-evident. The synonyms used, in turn, go back to their respective cosine similarities to the word "inflation".

Our corpus includes a significant number of documents (11%) that discuss inflation without explicitly using the term. To capture these documents, we developed an extensive pattern search incorporating common synonyms for inflation. Rather than relying on arbitrary word lists from synonym databases, we identified words used similarly to inflation within our dataset, based on their cosine similarity to the target word. To calculate word similarities, we adapted a procedure proposed by Moody (2017). First, we draw a random sub-corpus of size 50 000 from our corpus, which we tokenized and cleaned of stop-words. Based on this sub-corpus, we computed the normalized skipgram probabilities of each possible word tuple. That is, for each word in the sub-corpus, we compute the probability with which it appears near every other word (within a word window of size 4). Knowing the probability of a word appearing at all, we can thus determine which words appear next to each other more often than would be expected when they were independent. These values are stored in a large matrix, which is afterwards decomposed using singular value decomposition (SVD), yielding word embeddings for each term. Based on these word vectors, we determine the cosine similarity of individual words and detect those words that are most similar to inflation. (See Appendix B for more details.) This approach has the advantage that it accounts for the nuances of financial and economic reporting in our selected newspapers and reflects the actual word distributions within our corpus. Therefore, the resulting embeddings are more accurate for our purposes than pre-trained

[2]The search terms (in German) correspond to inflation, devaluation of money and rising prices

word embeddings that were trained on generic datasets (e.g., Wikipedia).

While our search pattern was optimized for recall and includes synonyms of inflation and price increases, it does not explicitly capture deflation-related terms. This reflects the fact that stable or even declining prices are often perceived as benign in public discourse (apart from economists) and thus receive limited media attention. Empirical checks confirm that articles focusing on deflation without mentioning inflation (hence, articles we miss with our search term) account for less than 3 percent of all inflation- or deflation-related articles. Still, future refinements could explore a dual-indicator setup that explicitly accounts for deflationary language.

Following these preprocessing steps, the resulting issue-specific analysis corpus comprises 50 495 articles, 21 318 of which are from Handelsblatt, 14 670 from Süddeutsche Zeitung and 14 507 from Die Welt.

3.2. Decomposing Inflation Perception into dynamic topics

While the IPI captures the overall salience of inflation in the news, it does not reveal how inflation is thematically framed. To uncover dominant narratives and identify distinct thematic patterns within inflation-related reporting, we apply a topic modeling approach based on the RollingLDA algorithm (Rieger et al., 2021). In the following, we first motivate the use of RollingLDA instead of other dynamic LDA-type approahces by emphasizing their limitations.

Limitations of classical Topic Modeling

In its original form (Blei et al., 2003), LDA is well-suited for the identification of media frames (DiMaggio et al., 2013). Frame being an inherently static concept and LDA being a *static* method, they fit together well over limited time horizons and for thematically

limited text corpora. Over longer time horizons, however, the correspondence between research object and method is less obvious. After all, what we are interested in is detecting the evolution of thematic trends. Rieger et al. (2021) construct a *dynamic* version of LDA, called RollingLDA, that allows topic structures to change over time by modeling the fading of collective memory, as newer versions of narratives overwrite older ones. New data are fitted to a topic model that is calculated based on a rolling window of past observations.

Our RollingLDA-based approach addresses a number of problems with existing dynamic topic model variants, as e.g. TOT (Wang and McCallum, 2006), dDTM (Blei and Lafferty, 2006), and cDTM (Wang et al., 2008). First of all, the method is computationally less demanding than the modeling the entire corpus from scratch after each update. In addition, there is no information leakage from articles of later time points used for the modeling of articles from earlier time points. For example, if the co-occurrence of terms like *virus* and *pandemic* becomes dominant in recent periods (due to the Covid crisis), standard topic models that consider the full corpus may assign earlier mentions of *virus*—which in the past might have been referred to computer viruses—to the same health-related topic, thereby distorting historical topic structures. Through avoiding such information leakage, a RollingLDA based time series can also be used in forecasting setups in compliance with evaluation standards. Furthermore, the topic structure is preserved in such a way that the topics remain comparable over time, while still enabling gradual changes. In particular, this means that topic matching from independent subsequent runs of topic models is not required, as it is typically done in other studies (e.g., Anonymous, 2024; Greene et al., 2014; Maier et al., 2018) and which, for example, poses the problem that the topic structure is almost never preserved in those subsequent runs. For the application in empirical economics RollingLDA has great value, as it allows us to build consistent, updatable time series of our overall Inflation Perception Indicator (IPI) as well economically interpretable topic-based sub-indicators. RollingLDA lends itself to a broad variety of text-based applications yet to be developed.

RollingLDA: model initialization

We begin by preprocessing the IPI subcorpus. All words are converted to lowercase and German umlauts are resolved. Since it has been shown that stop words are not critical for the modeling itself (Schofield et al., 2017), we only delete a conservative selection of stop words that do not contribute to the generation of topics or that might involve noise which would impede the later interpretation. Since it has also been shown that stemming has no substantial benefit and is sometimes even harmful (Schofield and Mimno, 2016), we use the word token in their original form, which also facilitates the later interpretation.

Here, to start the rolling approach, we consider all the articles published between January 1, 2001 and December 31, 2005. Using a rather low threshold, we determine the vocabulary for this initial modeling: all the words that occur more than five times in this time interval are considered. This procedure removes the long tail of very infrequently occurring words that provide very little information. The result is a sub-corpus of 13 895 texts with an average of 239 tokens from a 42 834 words-vocabulary for the first modeling period. These texts from the first five-year-period are modeled using the LDAPrototype procedure (Rieger et al., 2022a). After a thorough content analysis, we chose the model's parameter as $K = 10$ topics and accordingly as Dirichlet parameters $\alpha = \eta = 1/K$, while the Gibbs sampler iterates 200 times over the dataset. For details on the LDAPrototype approach and the content analysis, we refer to Section A in the Appendix.

RollingLDA: model updating

In the second modeling step of the rolling approach, we consider the articles from the subsequent first month of 2006, i.e. the 193 articles published between January 1 and January 31, 2006. By applying the seeding procedure of RollingLDA, we model the topic assignments of these 193 articles. However, we only use the last three months as memory, i.e. we initialize the model with the 863 articles from October to December 2005. The vocabulary is extended by words that occur more than five times in the new 193 articles

and that were not included in the vocabulary before. Employing this procedure, we add 11 words in the first month of 2006. The topic assignments of the new articles are initialized randomly and the Gibbs sampler iterates over each of the new articles another 200 times, while the topic assignments of all articles acting as initializing memory remain constant.

We apply the model updating procedure, which we described for the first month of 2006, on a rolling basis for all subsequent months, so that we finally obtain assignments to the 10 topics for the entire analysis corpus with an average of 280 tokens over a vocabulary of 48 487 different words. The modeling of newly occurring articles is performed analogously.

While using LDAPrototype for initialization ensures the reliability of the method, the restriction to three months as memory opens the possibility for the appearance of new topics or the mutation of existing ones. This parameter can be varied. However, three months are intuitive from the point of view that the memory spans one quarter. A larger number of months, i.e. a longer memory, could lead to very inflexible models, a reduction to fewer months to more flexible, but also to rapidly changing and instable topics. To check the memory parameter, we recommend looking at the self-similarity of the topics over time. Since we allow topic structures to change, we have to make sure that topics remain stable over time to a degree that comparability is ensured, i.e., that they actually deal with similar content. Certain actors may change, new terms may be coined, some words may fade from vocabulary while others become fashionable; nonetheless, a topic should contain articles about similar issues over the entire time horizon.

4. Results: Inflation perception over time

To understand how attention to inflation evolved in the public sphere, we examine both the intensity and the thematic composition of inflation-related media coverage over time. This section presents empirical results derived from our Inflation Perception Indicator

(IPI) and its topic-level decomposition based on the RollingLDA framework.

4.1. Corpus analysis

Figure 1 illustrates the dynamic relationship between media attention towards inflation and actual inflation rates in Germany over the past two decades. The graph displays two key metrics: the share of inflation-related articles in the overall corpus (referred to as the Inflation Perception Indicator or IPI) and the monthly Consumer Price Index (CPI) inflation data (year-over-year). In the 2000s, an initial period of relative calm is visible, where inflation oscillates around 2% and inflation is only a minor issue of interest. This changes in 2008 when the boom preceding the Great Financial Crisis (GFC) drives up inflation rates above 3% annually. Inflation coverage reaches a local peak in July 2008, in parallel with actual price developments. After that, inflationary pressures recede in an environment of plummeting financial markets, turning media attention away from inflation. In 2011, somewhat similar developments occur: inflation and its coverage are both on the rise, until the sovereign debt crisis rattles the Eurozone to an extent that leads to a gradual and prolonged decline of inflation rates. This time, however, inflation coverage does not revert to low pre-GFC levels, but stays at somewhat elevated levels even as CPI inflation falls towards zero, a period where price dynamics and their coverage seem somewhat out of sync. The Covid-19 pandemic from 2020 prompts an unprecedented slump in economic activity and a subsequent elimination of inflationary pressures; media attention duly declines. The post-pandemic recovery causes inflation readings to shoot up unexpectedly which is causing considerable media coverage broadly in line with actual developments.

The IPI curve progression indeed resembles the predictions of communication theory laid out in the introduction: media attention is largely driven by actual developments, with distinct waves of coverage being kicked-off by key-events (inflation spikes), subsiding afterwards as public attentiveness is re-directed to other issues. However, there is an exception:

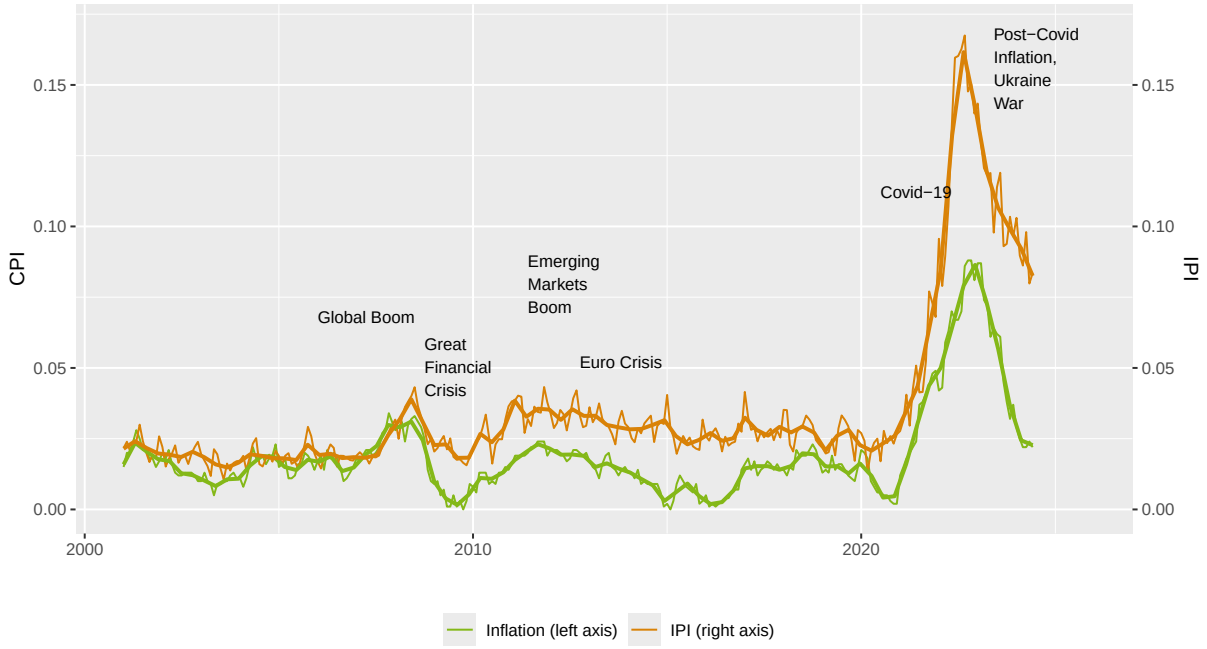

Figure 1: Inflation (CPI for Germany, y-o-y percentage change, green) and overall IPI (orange) are shown together with respective three month moving-averages. The IPI displays the share of the analysis corpus relative to the entire corpus. Source: authors' calculations, Deutsche Bundesbank

between 2012 and 2015 inflation coverage remains at elevated levels, even though inflation rates decline considerably. This effect can be linked to phase 5 in Downs' (1972) taxonomy, where issue attention stays above pre-cycle levels after a wave of coverage, a phenomenon that can be attributed to collective memory and the news value of consonance (e.g., Müller, 2023, Chapter 3), i.e., new information is incorporated into existing media narratives.

4.2. Topic analysis

Decomposing the "IPI corpus" into distinct topics as described in the previous section yields deeper insights into media focus. By applying the RollingLDA, we can identify ten clearly distinguishable topics, nine of which (comprising almost 90 per cent of the analysis corpus) are interpretable in terms of our research interest. Table 1 provides an overview of the model's topics. Labels were assigned manually based on the most characteristic words and a qualitative reading of representative documents. The table reflects the final

topic-word distribution, aggregating word counts across the entire observation period.

Table 1: Overview of the LDA model's topics, January 2001 - March 2024

Topic No.	Label	Share[a]	Topic content	Top words	Key events
1	Central Banks	14.1	Speculation about policy measures of advanced economies' central banks (rate changes, asset purchases/sales etc.)	ecb, fed, interest rate, central bank, inflation, interest rate hikes, monetary policy, central banks	7-08 (rate hike), 11-11 (rate drop), 1-15 (before QE), 7-19 (disc. of QE restart), 21/22 (inflation surprise)
2	News	12.5	News briefs covering data releases (CPI, market rates, oil price etc.)	percent, prices, gas, consumer, company, october, germany, year	inflation spikes (7-08, 2-11, 12-21), QE (1-15)
3	Emerging Markets	6.5	Inflationary developments in Argentina, Turkey, Venezuela, Iran...	government, country, Turkey, president, Russia, Argentina	EM booms (7-08, 1-11), Iran uprising (1-18), inflation spike (12-21)
4	Eurozone	12.4	Policy discussions and developments in other Eurozone countries and at EU level (fiscal stance, Stability and Growth Pact etc.)	china, germany, europe, must, eu, our, yes, usa, company, gas	Sov. Debt crisis (11-11), Greek stand-off (1-15), post- covid inflation surprise (10-21)
5	Private Investment	10.8	How to cope with low real yields, "private saver" perspective, focus on inflation hedges (Gold, Real Estate...)	banks, clients, tenants, interest, insurer, landlord, euro, says, bank, private	Boom (7-08), EZ uncertainty (sev. peaks 3-11 – 11-13), post-cov-19 inflation surprise (12-21)
6	Miscellaneous	11.4	Diverse		

[a] Share of analysis corpus (per cent)

Topic No.	Label	Share[a]	Topic content	Top words	Key events
7	Financial Markets	12.5	Financial Market developments and reactions to inflation risks	shares, investor, percent, dollar, dax, share prices, investors, interest, share, bonds	Nat. election (10-05), EM Booms (7-08, 3-11), Inaug. Trump (1-17), trade war (2-18), post-covid rally (3-21)
8	Companies	4.8	Developments in certain companies and sectors in Germany, focus on shareholder meetings and earnings calls, many calendar previews	million, corporation, company, turnover, euro, auris, airlines, customers, henkel, startups	Post-covid 19 inflation surprise (12-21)
9	German Politics	6.7	Collective bargaining, social, tax, fiscal policies – reactions to inflation	euro, scholz, spd, lindner, federal government, fdp, buergergeld, federal level, merz, gas price break	Inflation surprises in 7-08, 11-11 and 11-21
10	Raw Materials	8.9	Inflationary developments in Emerging Markets with particular focus on raw material demand and prices (gold, oil, copper, wheat...)	Ftx, yoon, dollar, Fisker, kim, oujoon, Thielemann, Bitcoin, mbc, tbs	Price hikes doe to trade tensions (most pronounced spike: 8-18), also inflation surprise 11-21

[a] Share of analysis corpus (per cent)

Descriptive analysis of the RollingLDA results reveal that some of these topics appear to correlate to a certain degree. This applies, for example, to Topic 5 (Private Investment) and Topic 7 (Financial Markets), which seems plausible given the shared set of vocabulary surrounding these topics (Figure 15, Appendix). Correlations are calculated based on the final topic-word-distribution "summarizing" all periods. The quarter-to-quarter similarity

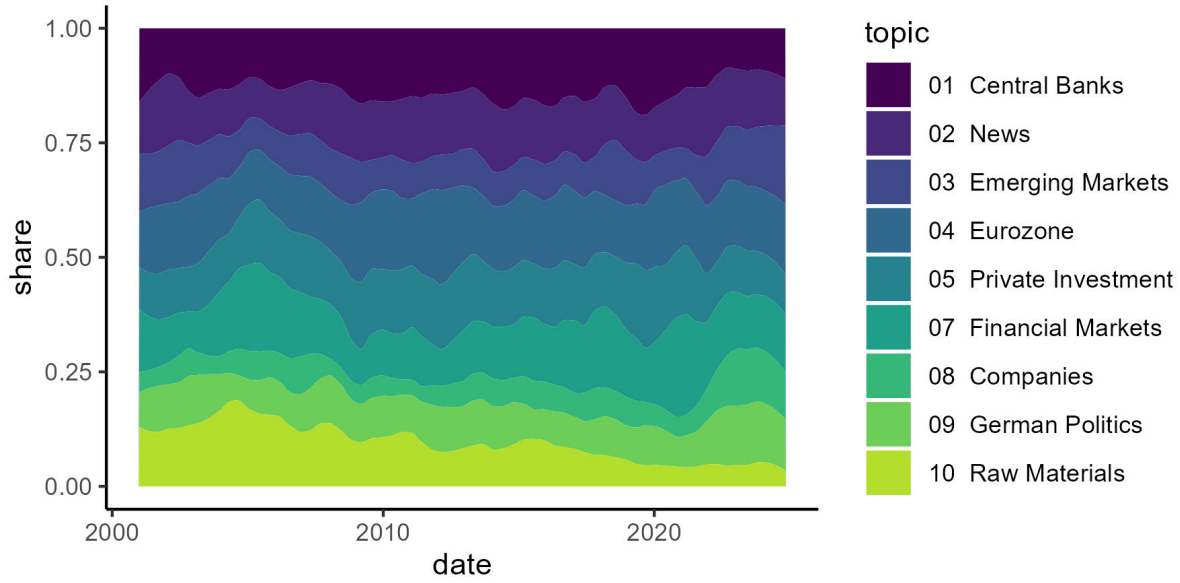

Figure 2: Relative contribution of individual topics to the Inflation Perception Indicator (IPI) over time. The graph shows smoothed monthly topic shares, stacked to sum to 1. Topic 6 (Miscellaneous) is excluded from the visualization for conceptual clarity, as it does not reflect a clearly interpretable aspect of inflation coverage.

of our $K = 10$ model (shown in Figure 14 in the Appendix) is rather high and stable for all the topics.

While Table 1 summarizes the semantic content of each topic in a static manner, media attention is highly dynamic and topics derived from a RollingLDA are typically modeled to represent thematic shifts during the observation period. To better understand how the salience of individual topics evolves over time, we track each topic's contribution to the Inflation Perception Indicator (IPI) on a monthly basis (Figure 2).

Some topics reflect relatively stable patterns across the observation window (e.g., *News*), while others are related to specific events. Notably, the topic *Financial Markets* (Topic 7) spikes around the global financial crisis in 2008, when German media intensely covered collapsing banks and investor panic. Towards the end of the sample, two other topics gain prominence: *German Politics* (Topic 9) and *Emerging Markets* (Topic 3), reflecting increasing attention to fiscal responses meant to buffer deteriorating real incomes and external inflation developments.

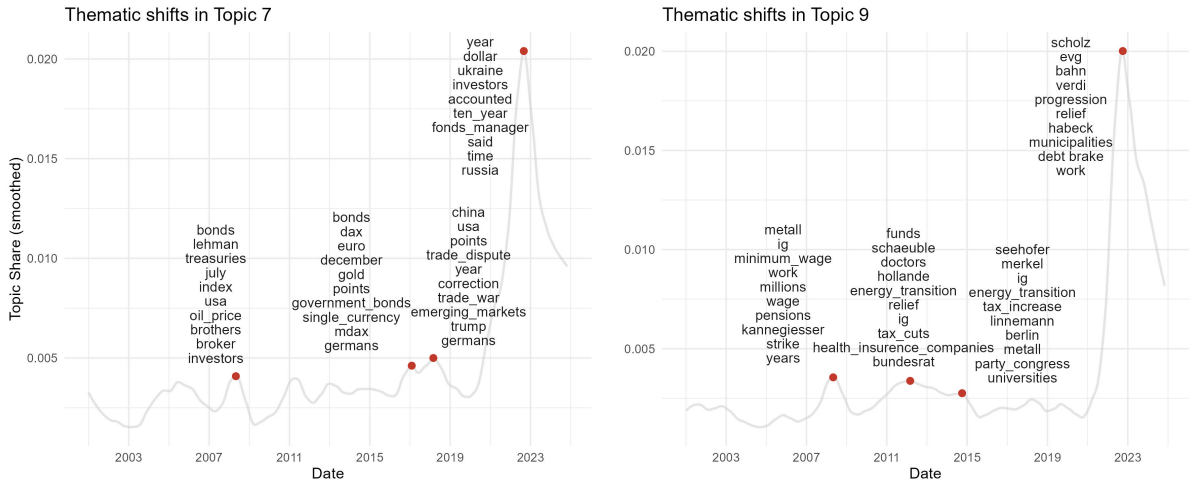

Figure 3: Temporal dynamics of Topic 7 (label: *Financial Markets*) and Topic 9 (label: *German Politics*) and indicative vocabulary shifts. The gray line shows the smoothed topic share over time. Red dots mark selected local maxima in topic salience. Word labels indicate terms with the highest *leave-one-out word impact*, that is, the greatest influence on the cosine distance between topic distributions three quarters apart.

To probe the evolution of individual topics more closely, we analyze semantic drifts within these topics over time. Figure 3 illustrates these dynamics for the topics *Financial Markets* and *German Politics*. The gray line shows the smoothed topic share over time, with red dots marking local salience peaks. For each peak, we identify the most characteristic words driving temporal change by computing their leave-one-out impact on the cosine similarity between topic-word distributions three quarters apart (Rieger et al., 2022b). The annotated terms—e.g., *lehman*, *trade_war*, *ukraine* in Topic 7—highlight major macro-financial shifts and illustrate how the semantic core of the topic adapts to unfolding events.

Comparing identical topics at different points in time, we observe that a large share of reporting on inflation is international or European in scope, in sum comprising about two thirds of the analysis corpus. Only the topics Private Investment, Companies and German Politics have a predominantly national focus. Given the international integration of product and financial markets and the European integration of monetary (and to some degree fiscal) policy, the breadth of coverage is hardly surprising. The Central Banks topic arguably is the most typically German one, in the sense that over large stretches

of time a rather conservative perspective on monetary policy prevails. Remarkably, the topic's salience remains considerable even during episodes of low and declining inflation during the 2010s, as newspapers kept questioning accommodating monetary polices. In characteristic articles, the ECB and its chiefs are the main protagonists; the Fed and other advanced-country central banks play minor roles. While the articles are mostly worded matter-of-factly, a certain bias towards a toughening of monetary policy can be detected, particularly when the ECB is concerned.[3]

Interestingly, some topics can be attributed to *exogenous* causes of inflation, while others capture *endogenous* responses to price dynamics. Since we are interested in potential causal relationships between inflation coverage and expectations formation, we focus on newspaper content about exogenous price developments, which we find in the topics Emerging Markets (topic 3) and Raw Materials (topic 10) in particular.[4] The latter is an obvious candidate, dealing with raw material prices that affect import prices directly, while the former stresses price dynamics in Emerging Markets that are exogenous from a German (i.e., intra-Euro area) perspective.

To be sure, there is a range of topics that contain exogenous as well as endogenous aspects of inflation dynamics. Central banks (topic 1), for instance, may react to exogenous changes in the price level, but their policies also cause price adjustments over the medium term. Eurozone governments (topic 4) may react to citizens' deteriorating purchasing power by subsidizing households; at the same time, loosening fiscal policy in response to an outbreak of inflation may well trigger second-round effects. A similar reasoning applies to (re-)distributional policies at the national level (topic 9). Some topics clearly focus on the effects of inflation on savers and financial markets (topics 5 and 7), but also

[3]Examples of headlines read: *"ECB: Rates to be raised if inflationary pressures ensue." "EZB: Zinsen werden bei Inflationsdruck erhöht", Handelsblatt, September 23, 2005; "Weber: EZB muß auf Inflationsgefahr reagieren", Welt, May 11, 2006; "Deutsche Inflation übersteigt EZB-Ziel", Handelsblatt, May 2, 2017; "Gefährliche Debatte um höhere Inflation", Welt, June 19, 2018; "Behutsame Abkehr von Anleihekäufen", Süddeutsche Zeitung, September 21, 2021; "EZB-Ratsmitglied warnt vor Inflation", Handelsblatt, December 7, 2021* "Weber: ECB must react to inflation risk." "German inflation above ECB target." "Dangerous debate about higher inflation." "Tentative shift away from bond purchases." "ECB council member warns of inflation."

[4]We use a similar idea as Caldara and Iacoviello (2022).

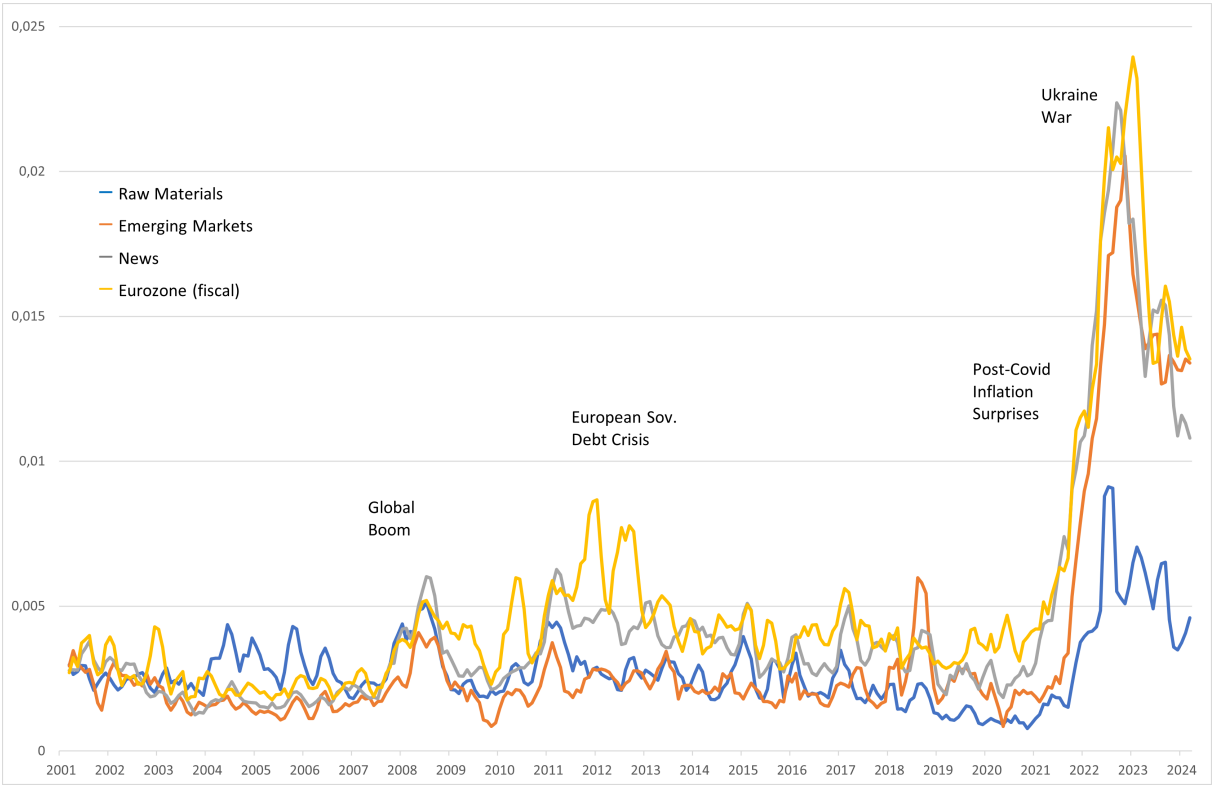

Figure 4: Selected IPI topics. The graphic displays the three-month moving averages. Values indicate the adjusted share of words that are assigned to each of the four topics. All values are weighted with the IPI to show their development in relation to the entire news coverage. Source: authors' calculations

on Companies (topic 8) that are predominantly portrayed as struggling with rising input prices, not as drivers of inflation by pushing up mark-ups. The topic News (topic 2) is exceptional in the sense that its content, short news pieces, refrains from insinuating causal relationships, but simply reports figures and forecasts.

In the subsequent econometric analysis, we include the two clear-cut exogenous topics, Emerging Markets and Raw Materials, as well as a control group comprising the topics News and Eurozone. Figure 4 shows that the four topics move largely in parallel. The situation starting in the second half of 2021 is unique in so far as topics News, Emerging Markets, and Eurozone shoot up to unprecedented levels simultaneously, while the volume of reporting on raw material prices remains considerably lower.

The topics' frequencies reveal that the heightened awareness of inflation in German media after the Great Financial Crisis of 2008 and the subsequent European Sovereign Debt

Crisis is driven in particular by an enduring preoccupation with fiscal policy and business-cycle developments across the Eurozone that get a lot more attention than during the days of the "great moderation" before the 2008 crash and the ensuing Euro crisis. During that period, raw material prices are only temporarily a minor source of concern, most notably in 2008, in 2018/19 following the deterioration of trade relations under the Trump administration, and in 2022 due to supply disruptions in the course of the Russian invasion of Ukraine.

To get a glimpse of the framing of the topics, one way is a qualitative analysis of the top articles (i.e., LDA-assigned texts characteristic of a topic's content):

> *Topic 3: Emerging Markets.* The topic comprises articles of inflation events in different Emerging Markets and developing economies such as Argentina, Turkey, Venezuela, and Iran. Top articles refrain from explicitly mentioning possible price spill-overs to Germany and the Euro area. However, implicit links to global inflation dynamics, that contain potentially useful information concerning domestic price developments, are frequently invoked. Typical headlines read: *"Zimbabwe abolishes its currency."* [5]*, "IMF supports Turkey's restructuring plans.", "A new president for Argentina."*
>
> *Topic 10: Raw Materials.* Top articles deal with price developments, but also with implications for exporting countries, such as Iran, Venezuela and Turkey. Hence, this topic partly overlaps with topic Emerging Markets. Characteristic headlines read: *"Boom in Asia clears oil markets."*[6]*, "Crop shortfalls ignite speculation.", "Oil pushes up Gold."*
>
> *Topic 4: Eurozone.* Fiscal policies of Eurozone countries, the recurring efforts to reform its framework (the stability and growth pact), and the enlargement of the

[5] "Zimbabwe schafft eigene Währung ab", Süddeutsche Zeitung, April 14, 2009, "Der IWF unterstützt türkischen Sanierungsplan", Handelsblatt, March 20, 2001; "Neuer Präsident in Argentinien", Welt, November 21, 2023

[6] "Asienboom fegt den Ölmarkt leer", Handelsblatt, September 29, 2004; "Ernteausfälle heizen Spekulation an", Handelsblatt, MKarch 7, 2006; "Öl treibt Gold", Handelsblatt, November 30, 2005

Eurozone are major themes of the top articles. The connection to inflation is rather indirect via potentially detrimental effects of high accumulated debt levels on the conduct of monetary policy. Before the recent bout of inflation, two events drove this topic's dynamics: the financial crisis of 2008 and the Eurozone sovereign debt crisis peaking in 2011. In recent years, fiscal reactions to the inflation outbreak have emerged as a central theme. Characteristic headlines read: *"Financial Crisis: Greek bankruptcy affects Germany, too."*[7], *"New Ideas for the Eurozone."*, *"Our Currency's stability is severely threatened."*

Topic 2: News. Short news pieces mostly related to the publication of new figures and forecasts concerning inflation and business-cycle developments. Articles mostly lack hints to causal relationships. The strictly factual nature of these articles is underlined by the top word in each and every month of the observation period, "prozent" (per cent). The topic contains new information about price developments at different levels of the economy (consumers, producers etc.) and in different countries. Declining as well as rising price levels are reported. The top articles suggest matter-of-factly, balanced reporting. Typical headlines read: *"Producer prices rise more slowly."*, *"Italy: Inflation rate drops slightly."*, *"Wholesale goods only slightly more expensive."*[8]

Summing up the perceived causes of inflation addressed in German coverage during the period considered in our analysis, we find that the blame is mainly put on the central bank. Fiscal policy plays its part, too, but to a lesser extent. Over the past two decades second-round effects in the form of rising wages and social benefits were not much of an issue in inflation reporting, though that has changed since higher CPI inflation induced higher wage demands.

[7] "Finanzkrise: Der Bankrott der Griechen streift auch Deutschland", Welt, December 10, 2009; "Neue Ideen für die Eurozone", Süddeutsche Zeitung, June 4, 2004; "Stabilität unserer Währung ist schwer gefährdet", Welt, July 9, 2005

[8] "Produzentenpreise steigen langsamer", Süddeutsche Zeitung, January 10, 2001, "Italien: Inflationsrate geht leicht zurück", Handelsblatt December 22, 2003; "Großhandelswaren nur etwas teurer", Süddeutsche Zeitung November 11, 2003

5. Impulse responses for news shocks in a threshold VAR model

In this section, we use a Bayesian threshold VAR model to analyze whether news shocks differ between topics on the one hand and between high and low inflation regimes on the other. To test whether the news topics actually contain different information about the inflation development, we estimate a BVAR model for each of the four news topics (Raw Materials, Emerging Markets, Eurozone and News) separately. The four topics cover inflation developments outside Germany, so that shocks in these topics can be considered exogenous. How the various inflation topics relate to the adjustment of private households inflation expectations is also analyzed in this section. As the literature on attention cycles suggests, news about the development of inflation only flows into the formation of expectations, if an attention threshold is exceeded. The attention threshold is defined on the basis of the inflation rate as in Korenok et al. (2023) and Pfäuti (2024). Moreover, variables of a standard New Keynesian model (production, inflation and interest rates) are included in the BVAR model.

The data for inflation expectations of private households are taken from the consumer survey of the European Commission. This survey is conducted monthly. In this survey, 2000 people are asked about their assessment about the current situation and expectations about future economic developments. With regard to future price developments, people are asked about their assessment of the development of consumer prices over the next twelve months compared to the last twelve months. Possible answers are “rise more sharply than before”, “rise to somewhat the same extent as before”, “rise to a lesser extent than before”, “remain more or less the same”, “decrease” and “do not know”. Based on these answers a monthly balance indicator is constructed.

In addition, we use the variables of the standard New Keynesian model. We therefore include the year-on-year growth rate of industrial production and the year-on-year change

of the CPI from the Federal Statistical Office. In addition, we use the one month Euribor from the Deutsche Bundesbank as the nominal interest rate.

We use the threshold BVAR model of Alessandri and Mumtaz (2017) to estimate two different regimes for the development of inflation, that is, we use the model

$$B_0 Y_t = \left[c_1 + \sum_{j=1}^{P} B_{1,j} Y_{t-j} + e_t \right] S_t + \left[c_2 + \sum_{j=1}^{P} B_{2,j} Y_{t-j} + e_t \right] (1 - S_t) \,, \tag{1}$$

where $S_t = 1 \Leftrightarrow z_{t-d} \leq z^*$ is the threshold variable. The vector of variables is $Y_t = \{y_t, \pi_t, r_t, \pi_t^e, ipi_{i,t}\}$ with industrial production y_t, inflation π_t, interest rate r_t, inflation expectations π_t^e and the respective ipi indicator $ipt_{i,t}$. Using this specification, the high inflation regime is defined as a period in which the actual inflation rate is above the threshold. The threshold is calculated from the lagged inflation rates.

The model is estimated using the Gibbs sampler with five steps. Since our model is very similar to Alessandri and Mumtaz (2017), we largely adopt their specification of the prior and the hyperparameters. In the first step, based on a given threshold, the two inflation regimes are defined. In step two and three, the coefficients of the BVAR models for the two regimes are estimated separately. Afterwards, the threshold is sampled in a random walk Metropolis step. In a final step, the delay parameter for the threshold is sampled. We use 3 500 iterations of the Gibbs sampler discarding the first 3 000 as burn-in. The number of lags for the VAR model in both regimes is specified using the AIC for a BVAR model without a threshold. The maximum number of delays for the threshold is set to 12 months.

The high and low inflation regimes identified within the estimation procedure depend on the number of lags in the BVAR model as well as on the optimal delay parameter for the threshold variable. In the estimation procedure, for both the News and Emerging Market topic, we use three lags. In this specification, mainly two high inflation periods are identified in Figure 5. Using four lags for the BVAR model, more high inflation periods

are identified for the Raw Material topic. The model for the Eurozone topic is estimated with two lags. Most of these periods are rather short, e.g., one or two months. The only exception is the year 2012 in which several months are classified as the high inflation regime. The average value for the threshold of inflation is 1.93%, which is very close to the estimate of 1.92% in Korenok et al. (2023) based on Google data.

To identify the structural shocks within this model we use zero restrictions based on the New Keynesian model as follows. Most important inflation expectations are a combination of current inflation and media news about future inflation based on the IPI. This relationship is used to impose zero restrictions in the fourth row of the B_0 matrix. Zero restrictions in row one to three are in line with the IS-curve, the Phillips curve and the Taylor rule. Moreover, we use media news topics that are related to foreign price developments or shocks like raw materials, emerging markets, Eurozone or news. We therefore assume that these news topic shocks are exogenous to the other variables of the system.

$$B_0 = \begin{bmatrix} 1 & b_{12} & b_{13} & 0 & b_{15} \\ b_{21} & 1 & 0 & b_{24} & 0 \\ b_{31} & b_{32} & 1 & 0 & b_{35} \\ 0 & b_{42} & 0 & 1 & b_{45} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

To identify the structural shock we use $u_t = B_t^{-1} e_t \sim (0, \Sigma_u)$ with $\Sigma_u = B_0^{-1} \Sigma_e B_0^{-1'}$(Kilian and Lütkepohl, 2017). Σ_e is a diagonal matrix with the variances on the main diagonal. To obtain comparable shocks, we set the variances in this matrix equal to one. To estimate the coefficients of the matrix B_0^{-1} we use the method-of-moments based on the second moment matrix Σ_u. As a robustness check we use a Cholesky decomposition with the news topic ordered last. In all cases median responses as well as 68% and 90% probability bands for a one unit shock are shown.

To test whether the impulse responses are different between the regimes, we calculate the

Figure 5: Inflation rate in the high and low inflation regime identified in the Emerging Market model January 2001 - March 2024. Shaded areas indicate the high inflation period identified by the model. Source: authors' calculations.

a posteriori distributions of the differences of the impulse responses. Median difference as well as the 68% and 90% probability intervals are shown.

Impulse responses for the four topics in the two attention regimes reveal some similarities, but also interesting differences. As can be seen in Figure 6, a shock to the Raw Material topic is followed by an increase in industrial production in both regimes. In this regard, the shock can be interpreted as a demand shock: more intensive reporting on dearer raw materials signals an uptick in global demand, which is associated with increasing manufacturing activity in the German manufacturing sector and accelerated price dynamics. Accordingly, we also find an increase in the interest rate in both regimes. The effect is significantly stronger in the high inflation regime. The effects on the inflation rate are only borderline significant at the 68% level. The effect in the low inflation regime is negative and in the high inflation regime slightly positive. An important distinction between the two regimes is that, in the high inflation regime, a news shock leads to a significant increase of consumer's inflation expectations, while there is no significant effect in the low

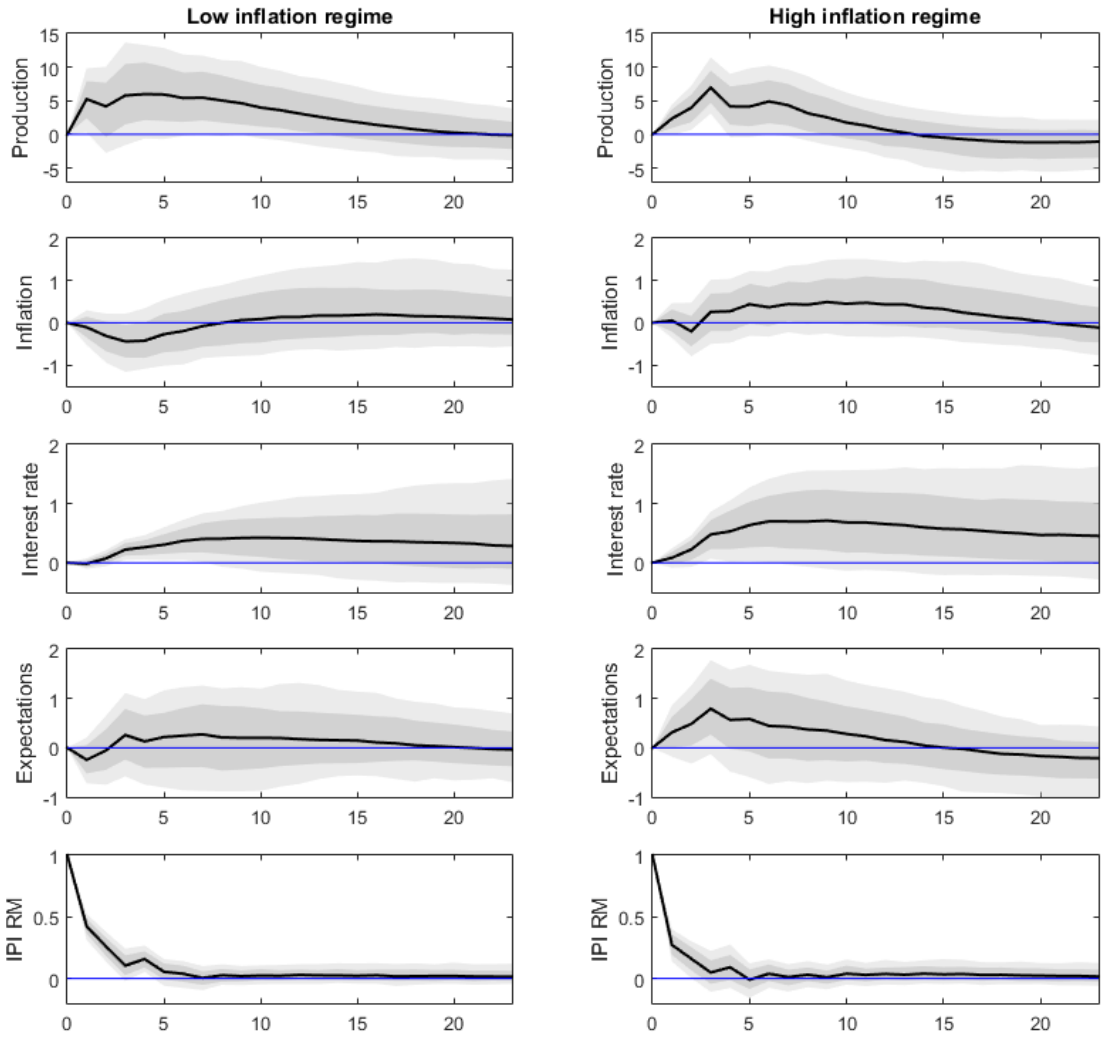

Figure 6: Impulse responses in a high and low inflation regime. The graphic displays impulse responses for a shock to the Raw Materials topic for two regimes. The five variable BVAR model was estimated with 4 lags. The threshold variable was constructed with a delay of 4 periods. We show the median responses as well as 68% probability bands (dark shaded areas) and 90% probability bands (light-shaded area). Source: authors' calculations.

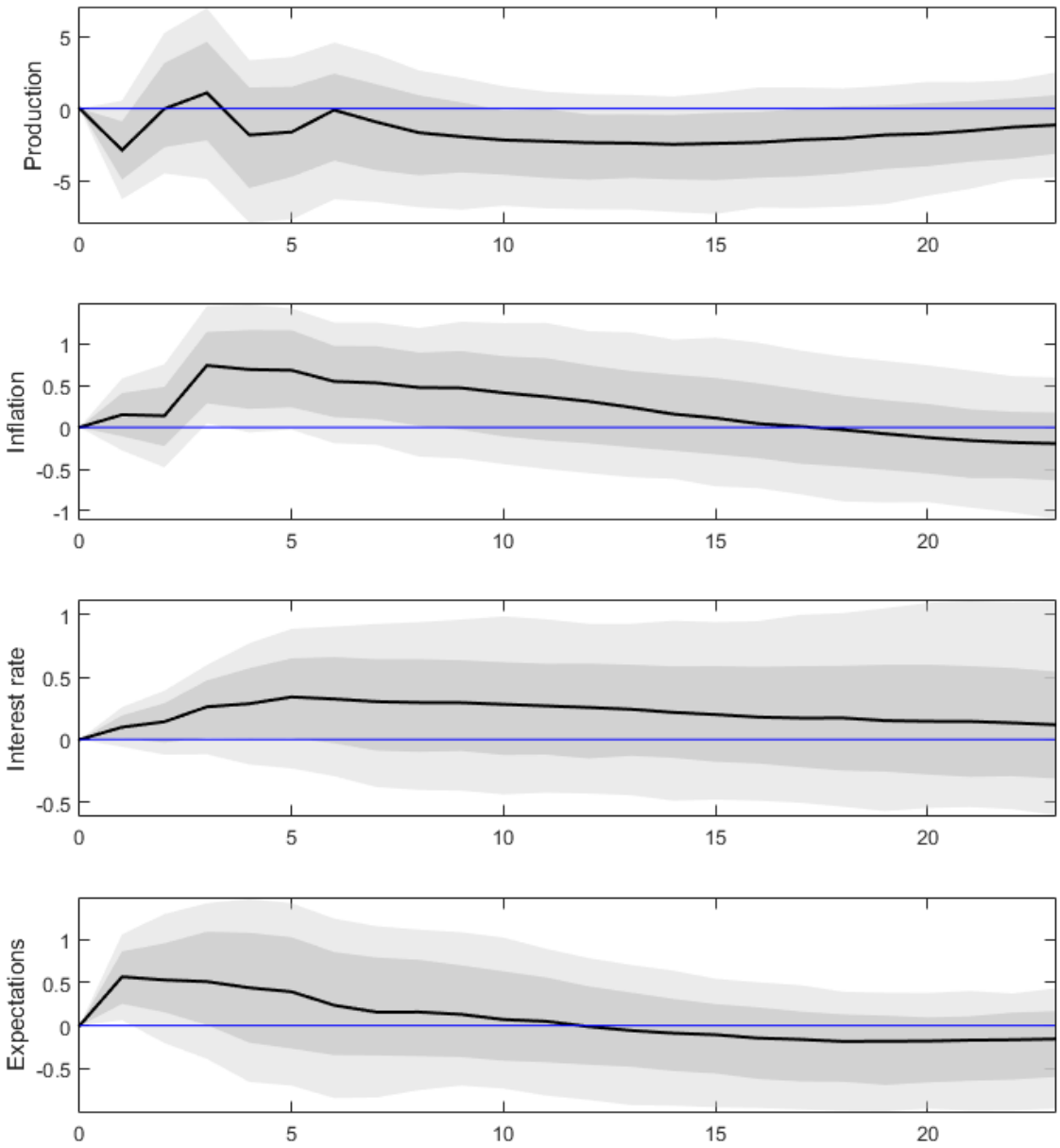

Figure 7: Differences between high and low inflation regime. The graphic displays the difference between impulse responses for a shock to the Raw Materials topic for two regimes. We show the median difference as well as 68% probability bands (dark shaded areas) and 90% probability bands (light-shaded area) of responses. Source: authors' calculations.

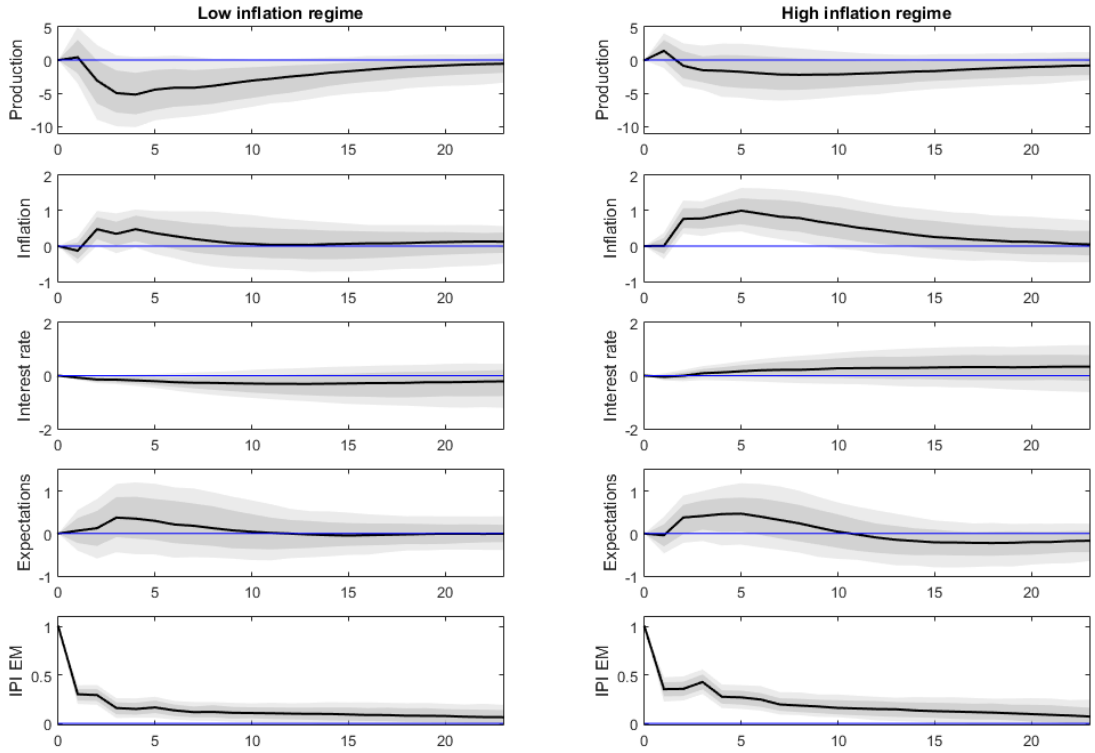

Figure 8: Impulse responses in a high and low inflation environment. The graphic displays impulse responses for a shock to the Emerging Markets topic for two regimes. The five variable BVAR model was estimated with 3 lags. The threshold variable was constructed with a delay of 3 periods. We show the median responses as well as 68% probability bands (dark shaded areas) and 90% probability bands (light-shaded area). Source: authors' calculations.

inflation regime. However, compared to the results of Pfäuti (2024), the effects are very short-lived.

Moreover, the differences between the regimes are significant with regard to household expectations and interest rates (Figure 7). In both cases there is a slightly positive difference between the high inflation and the low inflation regime. For expectations after one period and for inflation after three periods. For production we find a negative difference after two periods after the shock which is significant only at the 68% level. The difference for the interest rate response is not significant between the regimes.

According to Figure 8, a shock to the Emerging Markets topic leads to a decrease of industrial production in the low inflation regime, while there is no significant effect in the

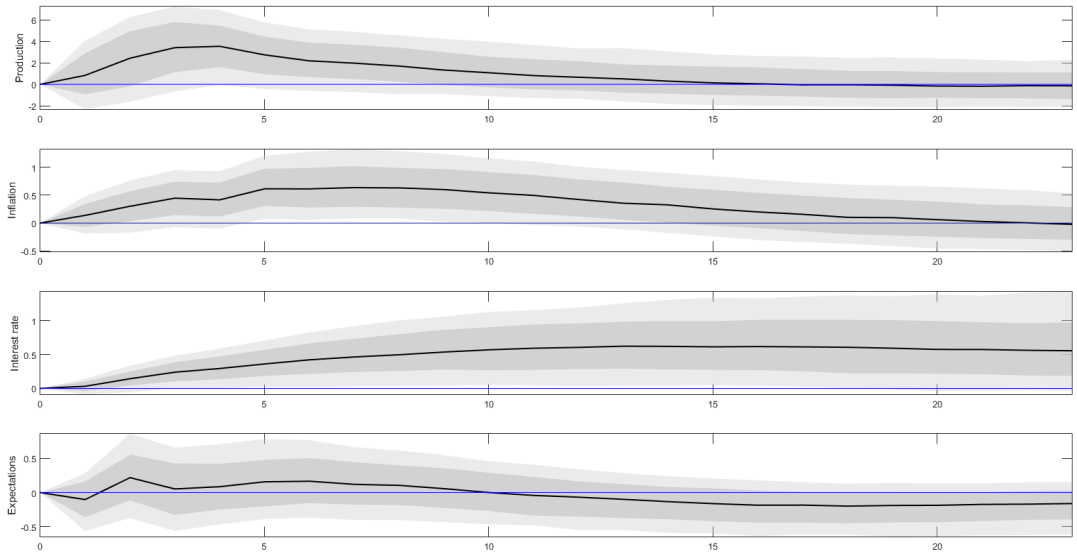

Figure 9: Differences between high and low inflation regime. The graphic displays the difference between impulse responses for a shock to the Emerging Markets topic for two regimes. We show the median difference as well as 68% probability bands (dark shaded areas) and 90% probability bands (light-shaded area) of responses. Source: authors' calculations.

high inflation regime. In addition, we find a small positive effect on the inflation rate in the low inflation regime and a strong positive effect in the high inflation regime as a supply shock. The interest rate fall in the low inflation regime and raise in the high inflation regime. However, both effects are only borderline significant. Interestingly, we find a significant positive effect on inflation expectations in the high inflation regime at the 68% level, while no significant effect is observed in the low inflation regime. Hence, more intensive reporting on inflation in emerging economies is associated with rising import prices of intermediate inputs, thereby hurting the supply-side of the domestic economy. In a high-inflation environment, where economic agents pay close attention to price developments, this leads to an uptick in domestic inflation, as producers pass on input price increases to consumers. Accordingly, interest rates are adjusted upwards. In a low-inflation environment, these effects are negligible, as in these circumstances firms are less inclined to pass on cost-increases.

Even in the case of the emerging market news shock, the effects are significantly stronger

in the high-inflation regime than in the low-inflation regime, as measured by the impulse response sequences (Figure 9. Moreover, we find significantly higher interest rate responses in the high inflation regime. The difference in the response of the real economy is only significant at the 68% level. The difference of expectations is not significant. One explanation could be that the monetary policy response prevented an increase in inflation expectations.

As displayed in Figure 10, shocks of the News topic generate impulse responses that are different from the two topics before. Most important in the high inflation regime, shocks do not have a significant effect on inflation expectations, while the effects in the low-inflation regime are negative. Moreover, in the high-inflation regime, there are positive effects on the interest rate and inflation, but there is no significant effect on industrial production. In the low inflation regime, there are positive effects on inflation and production. As we have seen in Section 4, these News topic shocks contain reports on current price trends and therefore differ from the patterns of macroeconomic supply and demand shocks.

When looking at the differences in impulse response sequences (Figure 11), significant differences in inflation are once again apparent, with a much stronger reaction in the high inflation regime. The differences in production and expectations are only significant at the 68% level. Production declines more sharply, while inflation expectations are slightly higher. There are no differences in interest rates.

Shocks to the Euro topic shown in Figure 12 mainly lead to significant effects in the low inflation regime. We find an increase in production and inflation at the 68% significant level. And a slight decline in the interest rate. In the high inflation regime, we find a small decline in the interest rate and inflation expectations. This decline in expectations is also visible in the difference between the two inflation regimes (Figure 13.

Taken together, the results confirm that news topics contain specific information about the inflation development like sources and possible effects on the economy and likely reactions

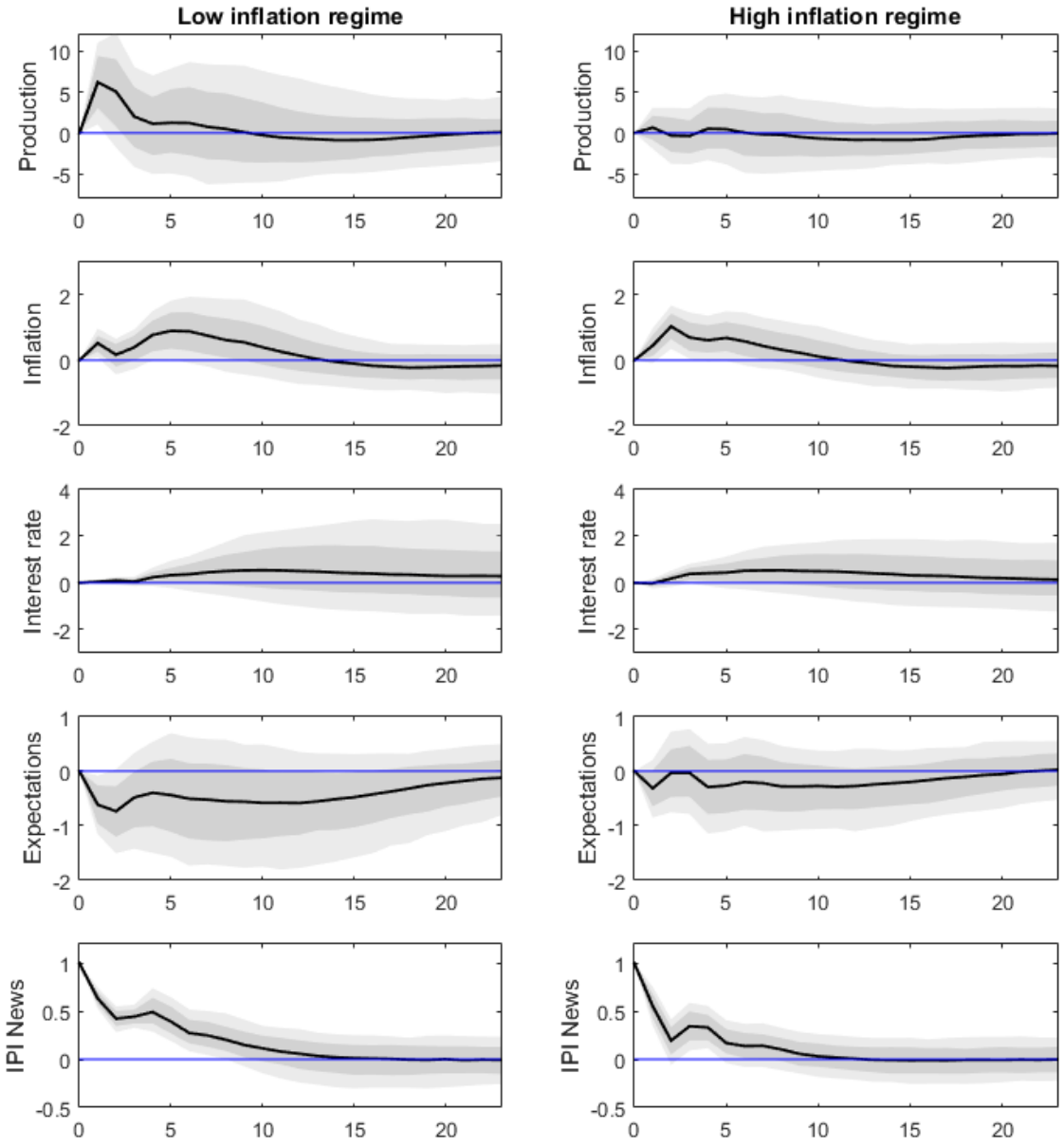

Figure 10: Impuls-responses in a high and low inflation environment. The graphic displays impuls-responses for a shock to the News topic for two regimes. The five variable BVAR model was estimated with 3 lags. The threshold variable was constructed with a delay of 3 periods. We show the median responses as well as 68% probability bands (dark shaded areas) and 90% probability bands (light-shaded area). Source: authors' calculations.

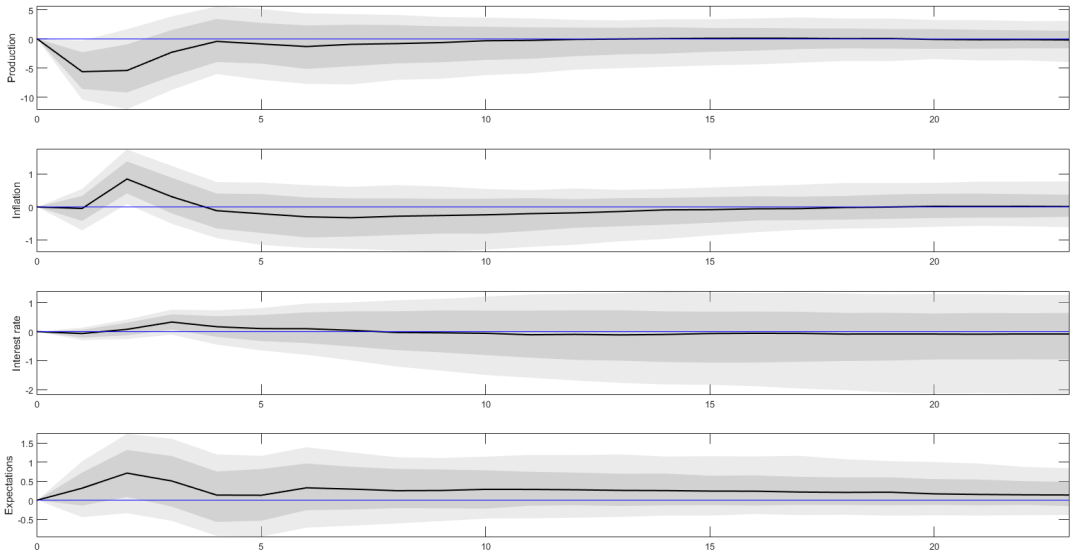

Figure 11: Differences between high and low inflation regime. The graphic displays the difference between impulse responses for a shock to the News topic for two regimes. We show the median difference as well as 68% probability bands (dark shaded areas) and 90% probability bands (light-shaded area) of responses. Source: authors' calculations.

of monetary policy. This explains why media coverage of inflation affects households' inflation expectations in different ways.

6. Conclusions

This paper examines how public narratives around inflation emerge and evolve, and whether these narratives matter for the formation of inflation expectations. We draw on the previously established Inflation Perception Indicator (IPI) to track the salience of inflation in public discourse over time, based on 50,000 articles from three major German newspapers. Using RollingLDA, a dynamic topic modeling approach, we decompose the IPI into interpretable sub-indicators that capture changing media narratives in a time-consistent manner.

A threshold VAR model is used to investigate whether news shocks have different effects

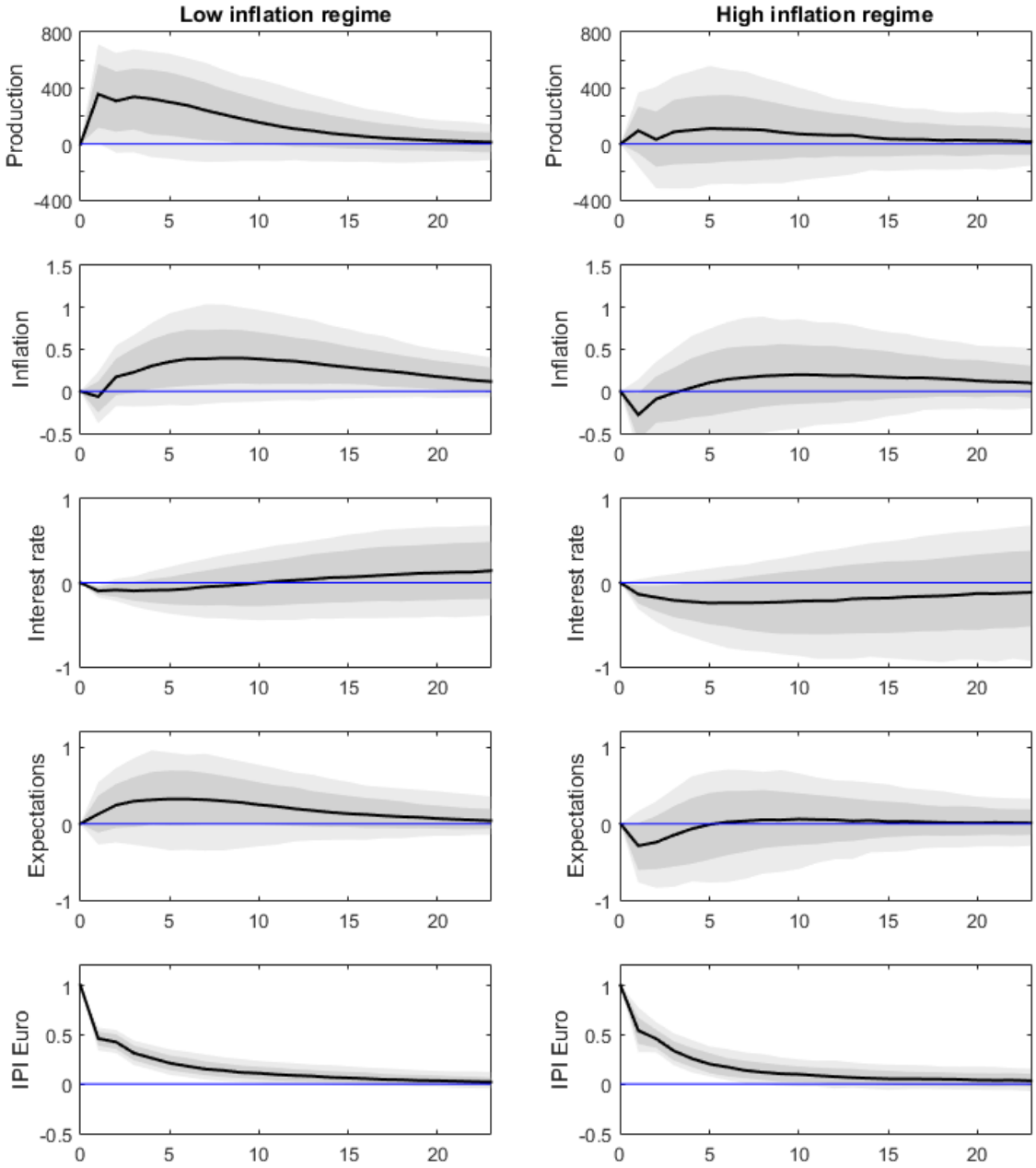

Figure 12: Impulse responses in a high and low inflation environment. The graphic displays impulse responses for a shock to the Eurozone topic for two regimes. The BVAR model was estimated with 2 lags. The threshold variable was constructed with a delay of 1 periods. We show the median responses as well as 68% probability bands (dark shaded areas) and 90% probability bands (light-shaded area). Source: authors' calculations.

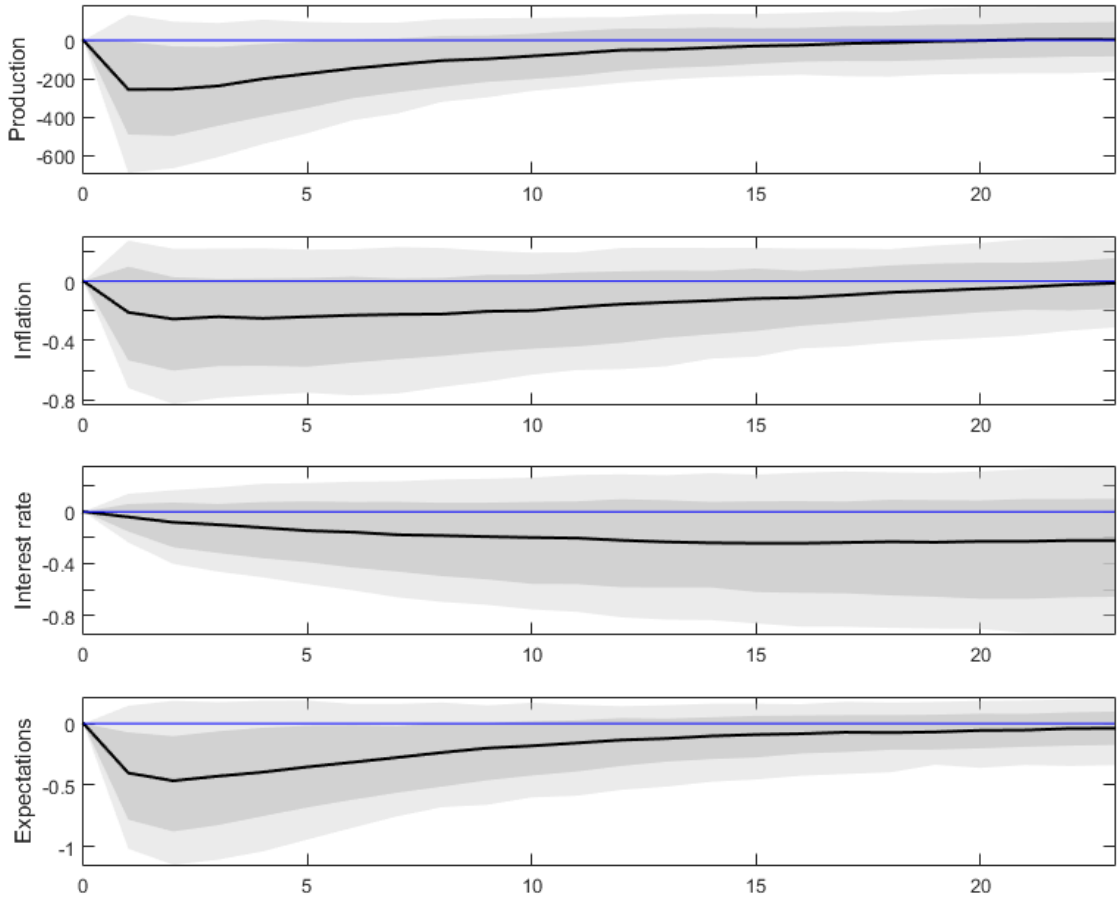

Figure 13: Differences between high and low inflation regime. The graphic displays the difference between impulse responses for a shock to the Eurozone topic for two regimes. We show the median difference as well as 68% probability bands (dark shaded areas) and 90% probability bands (light-shaded area) of responses. Source: authors' calculations.

on macroeconomic variables in different inflation regimes. The results show that news shocks associated with raw material prices, emerging economies and pure reporting of new statistical information differ depending on the level of inflation. In a high-inflation regime, people and firms pay close attention to price changes and adjust their behavior accordingly. In a low-inflation regime, these effects are weak or non-existent. This is in contrast to the Eurozone topic, where we cannot distinguish between different regimes.

Moreover, we show that the topics contain additional information about the type of real shocks. The impulse response sequences for news shocks associated with rising raw material prices can be interpreted as demand shocks, as dearer commodities signal an upswing of the global business cycle. Contrastingly, news shocks emanating from rising inflation in emerging economies show characteristics of a supply shock, as they are associated with rising costs of intermediate goods. In a high-inflation environment, firms are passing these import price increases on to consumers, while, in a low-inflation regime, this effect is rather subdued.

We include the inflation expectations of private households in the analysis as a possible channel that can explain the different size of the effects of news shocks. Recent empirical studies argue that private households pay more attention to inflation developments, when these exceed a certain threshold value. We also find this effect in our analysis. Furthermore, we show that higher inflation is associated with greater media coverage. The higher intensity of reporting is one possible channel that lead to adjustments of inflation expectations. In addition, it is obvious that different topics are related to the formation of expectations of economic actors.

Hence, according to our results, it is not just the gross intensity of coverage that shapes expectations, but more so the subtler facets of reporting. The effects on inflation perception critically hinge on the specific issues and frames highlighted by the media. Our findings show different response patterns for different kinds of inflation shocks, which the IPI methodology ably captures. In particular, we find that shocks related to Raw

Material prices show effects like a demand shock, while news shocks spilling over from abroad resemble the pattern of supply shocks. Still, the sheer volume of public attention matters. In an environment of intensive inflation reporting the effects on expectations are decisively stronger than in a low inflation regime.

These findings underline the critical importance of monetary policy in managing inflation expectations in a setting of heightened inflation attention. When operating in a low inflation regime, the economy's tolerance to glitches in policy making is rather benign. In contrast, the central bank's role is more intricate when the public pays close attention to price level movements. In these circumstances, monetary policy needs to be decisive and cautious at the same time. Arguably, temporarily erring on the side of tightness is less costly than risking the unmooring of expectations.

As the media have the tendency to lower reporting intensity once an issue loses its newness, the resulting inattention is set to induce households to adapt their expectations in a delayed manner, even if actual inflation rates fall successively. This type of interaction between the media and economic agents is likely to contribute to the stickiness of inflation expectations, complicating the job of policy makers. The gap between inflation perception and actual price developments may be aggravated, if a convincing "inflation narrative" emerges that puts the blame on domestic institutions assigned to keep inflation under control. In fact, the results of the IPI model can be interpreted as manifestations of a German inflation narrative that attributes the role of villains to central banks, finance ministries, and raw materials-exporting countries (Müller et al., 2022).

However, our findings come with a crucial caveat. Most of the observation period was characterized by an extensive phase of more or less stable price levels. As inflation remained at low levels for many years, it comes as little surprise that we find low levels of attention – and insignificant links to expectation formation – for much of the observation period. All this may change, though, were the world economy to enter a prolonged high inflation regime (Borio et al., 2023). In such radically altered circumstances, the behavior

of individuals, company executives and, not least, the media is likely to change profoundly. If and how these changes come about will be an issue worthwhile to be studied in future research.

Acknowledgments

This work was financially supported by the Mercator Research Center Ruhr (MERCUR) with project number Ko-2022-0015. In addition, the authors gratefully acknowledge the computing time provided on the Linux HPC cluster at TU Dortmund University (LiDO3), partially funded in the course of the Large-Scale Equipment Initiative by the German Research Foundation (DFG) as project 271512359

References

Alessandri, Piergiorgio and Haroon Mumtaz (2017). "Financial conditions and density forecasts for US output and inflation". In: *Review of Economic Dynamics* 24, pp. 66–78. ISSN: 1094-2025. DOI: https://doi.org/10.1016/j.red.2017.01.003. URL: https://www.sciencedirect.com/science/article/pii/S1094202517300042.

Angelico, Cristina, Juri Marcucci, Marcello Miccoli, and Filippo Quarta (2022). "Can we measure inflation expectations using Twitter?" In: *Journal of Econometrics* 228.2, pp. 259–277. DOI: 10.1016/j.jeconom.2021.12.008.

Anonymous (2024). "BERTrend: Neural Topic Modeling for Emerging Trends Detection". In: *Submitted to ACL Rolling Review - June 2024*. under review. URL: https://openreview.net/forum?id=TwJ2q0W0OH.

Berger, Helge, Michael Ehrmann, and Marcel Fratzscher (2011). "Monetary policy in the media". In: *Journal of Money, Credit and Banking* 43.4, pp. 689–709. DOI: 10.1111/j.1538-4616.2011.00392.x.

Blei, David M. and John D. Lafferty (2006). "Dynamic topic models". In: *Proceedings of the 23rd International Conference on Machine Learning*. Association for Computing Machinery, pp. 113–120. DOI: 10.1145/1143844.1143859.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003). "Latent Dirichlet Allocation". In: *Journal of Machine Learning Research* 3, pp. 993–1022. ISSN: 1532-4435. URL: https://jmlr.csail.mit.edu/papers/v3/blei03a.html.

Borio, Claudio E. V., Marco Lombardi, James Yetman, and Egon Zakrajšek (2023). "The two-regime view of inflation". In: *BIS Papers* 133.

Bracha, Anat and Jenny Tang (2022). "Inflation levels and (in)attention". In: *FRB of Boston Working Paper* 22-4. DOI: 10.29412/res.wp.2022.04.

Caldara, Dario and Matteo Iacoviello (2022). "Measuring Geopolitical Risk". In: *American Economic Review* 112.4.

Chang, Jonathan, Jordan Boyd-Graber, Sean Gerrish, Chong Wang, and David M. Blei (2009). "Reading Tea Leaves: How Humans Interpret Topic Models". In: *NIPS: Advances in Neural Information Processing Systems*. Ed. by Y. Bengio, D. Schuurmans, J.

Lafferty, C. Williams, and A. Culotta. Vol. 22. Curran Associates Inc., pp. 288–296. URL: https://papers.nips.cc/paper/2009/hash/f92586a25bb3145facd64ab20fd554ff-Abstract.html.

Coibion, Olivier and Yuriy Gorodnichenko (2015). "Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation". In: *American Economic Journal: Macroeconomics* 7.1, pp. 197–232. DOI: 10.1257/mac.20130306.

Conrad, Christian, Zeno Enders, and Alexander Glas (2022). "The role of information and experience for households' inflation expectations". In: *European Economic Review* 143, p. 104015. DOI: 10.1016/j.euroecorev.2021.104015.

D'Acunto, Francesco, Ulrike Malmendier, and Michael Weber (2023). "What do the data tell us about inflation expectations?" In: *Handbook of Economic Expectations*. Ed. by Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw. Academic Press, pp. 133–161. DOI: 10.1016/B978-0-12-822927-9.00012-4.

DiMaggio, Paul, Manish Nag, and David Blei (2013). "Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding". In: *Poetics* 41.6, pp. 570–606. DOI: 10.1016/j.poetic.2013.08.004.

Downs, Anthony (1972). "Up and down with ecology: The issue attention cycle". In: *The Public Interest* 28, pp. 38–50.

Dräger, Lena and Giang Nghiem (2021). "Are consumers' spending decisions in line with a euler equation?" In: *Review of Economics and Statistics* 103.3, pp. 580–596. DOI: 10.1162/rest_a_00909.

Flynn, Joel P. and Karthik Sastry (2024). "Attention Cycles". In: *Working Paper*. URL: 10.2139/ssrn.3592107.

Greene, Derek, Derek O'Callaghan, and Pádraig Cunningham (2014). "How Many Topics? Stability Analysis for Topic Models". In: *Machine Learning and Knowledge Discovery in Databases*. Springer, pp. 498–513. DOI: 10.1007/978-3-662-44848-9_32.

Griffiths, Thomas L. and Mark Steyvers (2004). "Finding scientific topics". In: *Proceedings of the National Academy of Sciences* 101.suppl 1, pp. 5228–5235. ISSN: 0027-8424. DOI:

10.1073/pnas.0307752101. URL: http://www.pnas.org/content/101/suppl_1/5228.

Hansen, Stephen, Michael McMahon, and Matthew Tong (2019). "The long-run information effect of central bank communication". In: *Journal of Monetary Economics* 108, pp. 185–202. DOI: 10.1016/j.jmoneco.2019.09.002.

Hoyle, Alexander, Pranav Goel, Andrew Hian-Cheong, Denis Peskov, Jordan Boyd-Graber, and Philip Resnik (2021). "Is Automated Topic Model Evaluation Broken? The Incoherence of Coherence". In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan. Vol. 34. Curran Associates, Inc., pp. 2018–2033. URL: https://proceedings.neurips.cc/paper_files/paper/2021/file/0f83556a305d789b1d71815e8ea4f4b0-Paper.pdf.

Kilian, Lutz and Helmut Lütkepohl (2017). "Estimation Subject to Short-Run Restrictions". In: *Structural Vector Autoregressive Analysis*. Themes in Modern Econometrics. Cambridge University Press, pp. 241–268.

Korenok, Oleg, David Munro, and Jiayi Chen (2023). "Inflation and Attention Thresholds". In: *The Review of Economics and Statistics*, pp. 1–28. DOI: 10.1162/rest_a_01402.

Lamla, Michael J and Sarah M Lein (2015). "Information rigidities, inflation perceptions, and the media: Lessons from the Euro Cash Changeover". In: *Economic Inquiry* 53.1, pp. 9–22. DOI: 10.1111/ecin.12121.

Larsen, Vegard H, Leif Anders Thorsrud, and Julia Zhulanova (2021). "News-driven inflation expectations and information rigidities". In: *Journal of Monetary Economics* 117, pp. 507–520. DOI: 10.1016/j.jmoneco.2020.03.004.

Lörcher, Ines and Irene Neverla (2015). "The dynamics of issue attention in online communication on climate change". In: *Media and Communication* 3.1, pp. 17–33. DOI: 10.17645/mac.v3i1.253.

Maier, Daniel, A. Waldherr, P. Miltner, G. Wiedemann, A. Niekler, A. Keinert, B. Pfetsch, G. Heyer, U. Reber, T. Häussler, H. Schmid-Petri, and S. Adam (2018). "Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodol-

ogy". In: *Communication Methods and Measures* 12.2-3, pp. 93–118. DOI: 10.1080/19312458.2018.1430754.

Moody, Chris (2017). *Stop using word2vec*. Online, accessed: 5 April 2024. URL: https://multithreaded.stitchfix.com/blog/2017/10/18/stop-using-word2vec/.

Müller, Henrik (2023). *Challenging economic journalism: Covering business and politics in an age of uncertainty*. 1st ed. Palgrave Macmillan Cham. DOI: 10.1007/978-3-031-31030-0.

Müller, Henrik, Tobias Schmidt, Jonas Rieger, Lena Maria Hufnagel, and Nico Hornig (2022). "A German Inflation Narrative". In: *DoCMA Working Paper* 9. DOI: 10.17877/de290r-22632.

Nimark, Kristoffer P and Stefan Pitschner (2019). "News media and delegated information choice". In: *Journal of Economic Theory* 181, pp. 160–196. DOI: 10.1016/j.jet.2019.02.001.

Pfäuti, Oliver (2024). "The Inflation Attention Threshold and Inflation Surges". In: *Working Paper*.

Reis, Ricardo (2006). "Inattentive producers". In: *The Review of Economic Studies* 73.3, pp. 793–821. ISSN: 0034-6527. DOI: 10.1111/j.1467-937X.2006.00396.x.

Rieger, Jonas (2020). "ldaPrototype: A method in R to get a Prototype of multiple Latent Dirichlet Allocations". In: *Journal of Open Source Software* 5.51, p. 2181. DOI: 10.21105/joss.02181.

Rieger, Jonas, Nico Hornig, Tobias Schmidt, and Henrik Müller (2023). "Early warning systems? Building time consistent perception indicators for economic uncertainty and inflation using efficient dynamic modeling". In: *Proceedings of the 3rd Workshop on Modelling Uncertainty in the Financial World*. URL: https://github.com/JonasRieger/mufin23.

Rieger, Jonas, Carsten Jentsch, and Jörg Rahnenführer (2021). "RollingLDA: An update algorithm of Latent Dirichlet Allocation to construct consistent time series from textual data". In: *Findings of the Association for Computational Linguistics: EMNLP 2021*.

Association for Computational Linguistics, pp. 2337–2347. DOI: 10.18653/v1/2021.findings-emnlp.201.

Rieger, Jonas, Carsten Jentsch, and Jörg Rahnenführer (2022a). "LDAPrototype: A model selection algorithm to improve reliability of latent Dirichlet allocation." In: *Preprint available at Research Square*. DOI: 10.21203/rs.3.rs-1486359/v1.

Rieger, Jonas, Kai-Robin Lange, Jonathan Flossdorf, and Carsten Jentsch (2022b). "Dynamic change detection in topics based on rolling LDAs." In: *Text2Story@ ECIR*, pp. 5–13.

Schofield, Alexandra, Måns Magnusson, and David Mimno (2017). "Pulling Out the Stops: Rethinking Stopword Removal for Topic Models". In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Ed. by Mirella Lapata, Phil Blunsom, and Alexander Koller. Association for Computational Linguistics, pp. 432–436. URL: https://aclanthology.org/E17-2069.

Schofield, Alexandra and David Mimno (2016). "Comparing Apples to Apple: The Effects of Stemmers on Topic Models". In: *Transactions of the Association for Computational Linguistics* 4. Ed. by Lillian Lee, Mark Johnson, and Kristina Toutanova, pp. 287–300. DOI: 10.1162/tacl_a_00099.

Sims, Christopher A (2003). "Implications of rational inattention". In: *Journal of Monetary Economics* 50.3, pp. 665–690. DOI: 10.1016/S0304-3932(03)00029-1.

Stryker, Jo Ellen, Ricardo J Wray, Robert C Hornik, and Itzik Yanovitzky (2006). "Validation of database search terms for content analysis: The case of cancer news coverage". In: *Journalism & Mass Communication Quarterly* 83.2, pp. 413–430. DOI: 10.1177/107769900608300212.

Ter Ellen, S, Vegard H Larsen, and Leif Anders Thorsrud (2020). *Narrative monetary policy surprises and the media: How central banks reach the general public*. [Online, accessed 5 April 2024]. URL: https://voxeu.org/article/narrative-monetary-policy-surprises-and-media.

Ter Ellen, Saskia, Vegard H Larsen, and Leif Anders Thorsrud (2022). "Narrative monetary policy surprises and the media". In: *Journal of Money, Credit and Banking* 54.5, pp. 1525–1549. DOI: 10.1111/jmcb.12868.

Waisbord, Silvio and Adrienne Russell (2020). "News flashpoints: Networked journalism and waves of coverage of social problems". In: *Journalism & Mass Communication Quarterly* 97.2, pp. 376–392. DOI: 10.1177/1077699020917116.

Waldherr, Annie (2014). "Emergence of news waves: A social simulation approach". In: *Journal of Communication* 64.5, pp. 852–873. DOI: 10.1111/jcom.12117.

Wang, Chong, David Blei, and David Heckerman (2008). "Continuous time dynamic topic models". In: *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence*. AUAI Press, pp. 579–586.

Wang, Xuerui and Andrew McCallum (2006). "Topics over time: a non-Markov continuous-time model of topical trends". In: *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, pp. 424–433. DOI: 10.1145/1150402.1150450.

Watanabe, Kohei and Alexander Baturo (2024). "Seeded Sequential LDA: A Semi-Supervised Algorithm for Topic-Specific Analysis of Sentences". In: *Social Science Computer Review* 42.1, pp. 224–248. DOI: 10.1177/08944393231178605.

Weber, Michael, Francesco D'Acunto, Yuriy Gorodnichenko, and Olivier Coibion (2022). "The subjective inflation expectations of households and firms: Measurement, determinants, and implications". In: *Journal of Economic Perspectives* 36.3, pp. 157–184. DOI: 10.1257/jep.36.3.157.

A. Topic modeling details

For the initialization of the RollingLDA model, we make use of the selection method LDAPrototype (Rieger et al., 2022a). The method resolves the problem of arbitrariness that plagues the original LDA method which produces fundamentally different models with each run due to the random initialization of the Gibbs sampler (Griffiths and Steyvers, 2004). This problem exists even if exactly identical data and parameter settings are applied, running counter to the scientific requirement of reproducibility. By solving the problem of arbitrary selection, the reliability of findings is improved. Prototyping is achieved following a typical statistical approach: for a given combination of parameters a number of models is calculated (usually about 100), from which the particular model is determined that is most similar to all the other models in the set. For this purpose, pairwise model similarities are calculated using a tailor-made similarity measure for LDA models. These similarities are determined measuring the deviation from strictly matched topics in resulting local clusters, which are created based on a hierarchical clustering result of topics using pairwise topic similarities of two LDA results considered. These deviations are computed for all possible pairs of LDA models. The LDA that has the highest average similarity to all others is selected as the prototypical LDA. The methodology is implemented in the corresponding R package ldaPrototype (Rieger, 2020).

In addition to the LDAPrototype method for initial estimates of the model, we make use of RollingLDA that uses preceding LDA results as an initialization for subsequent quarters and iterates the collapsed Gibbs sampler over the new data only. The topic assignments of all the previously modeled articles remain constant and we obtain assignments to the existing topics solely for all new articles. The process of fitting new data to a predefined topic model is known as "seeding". RollingLDA can therefore also be seen as a kind of rolling variant of a seeded LDA (Watanabe and Baturo, 2024). Compared to other existing dynamic topic models, such as TOT (Wang and McCallum, 2006), dDTM (Blei and Lafferty, 2006), and cDTM (Wang et al., 2008), RollingLDA does not suffer from information leakage and ensures that no information from chunks in later periods affects

the topics in earlier periods during modeling. Hence, RollingLDA is ideal for practical use for monitoring of textual data and allows to construct time series to be used in out-of-time forecasting scenarios, for which TOT, dDTM and cDTM are not suitable due to the global modeling of the topics.

LDA requires the choice of parameter K, that is, the number of topics the algorithm is set to produce. This is a critical part of the analysis. An inadequate value of K results in topics that are thematically indistinguishable and therefore not applicable to the research questions. Instead, K should be set at a value where topics are formed that are coherent in the sense that they are separately interpretable by human researchers according to their research questions. Setting K arbitrarily, or according to some mathematical optimization approach, runs the risk of producing irrelevant result (Chang et al., 2009; Hoyle et al., 2021). We produced four LDA models on the corpus, with K set to 6, 8, 10, and 12 respectively, and resorted to an "eyeballing" procedure: three coders labeled the topics of each model independently, making use of the most characteristic word ("top words") and articles ("top articles") as well as each topic's frequency distributions over time. A value of $K = 10$ was found to be the most appropriate in terms of our research interest.

B. Search for inflation synonyms

Figure 14: Quarter-to-quarter intra-topic cosine similarities (black; gray: month-to-month, red: quarter-to-first-quarter, blue: quarter-to-last-quarter).

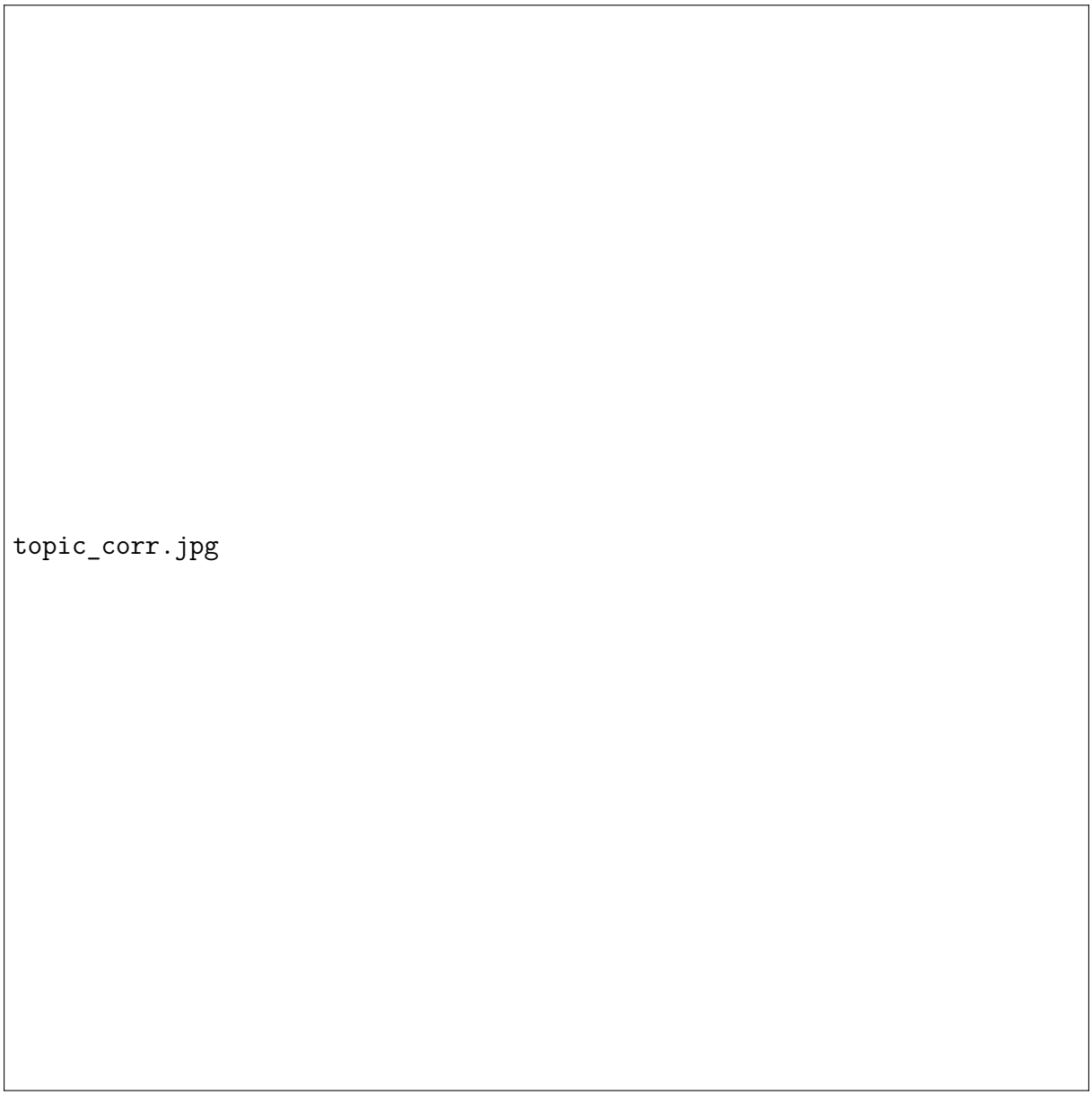

Figure 15: Topic correlations based on topic-word distribution across all periods.