The Role of Media Coverage in Shaping Household Inflation Expectations: An Analysis of ECB Press Conferences and News Reporting

Jasper Bär*

Abstract

PLACEHOLDER Abstract

Keywords: PLACEHOLDER Keyword 1, Keyword 2, Keyword 3, Keyword 4, Keyword 5

*PLACEHOLDER
Email address: `ja.baer@stat-econ.uni-kiel.de` (Jasper Bär)
URL: `PLACEHOLDER Github url`

1. Introduction

Central bank communication has become a vital instrument in modern monetary policy (Blinder et al., 2017). Central banks use communication as a tool for guiding inflation expectations and ensuring trust in their monetary policy. Historically, much of this communication has been directed towards financial experts. However, in recent years, central banks have increasingly been reaching out to the general public (Blinder et al., 2022). The media plays a critical role in this process by disseminating central bank communication to a broader audience. Therefore, it is essential to understand how the media reports on central bank communication to analyze the impact of central bank communication on the general public.

Since Carroll (2003) created a model for inflation expectations that considers the amount of media coverage, several studies have analyzed the role of media coverage for the inflation expectation forming process. Van der Cruijsen et al. (2015) find a small effect of media on inflation expectations and perception for Sweden. Similarly, Lamla & Lein (2014) find that media reporting influences the accuracy of German household inflation expectations. (Larsen et al., 2021) show that the news topics also have predictive power for inflation expectations. Ehrmann et al. (2017) show that increased media coverage leads to the strongest improvements in inflation expectations accuracy during recessions and for individuals with pessimistic views or financial difficulties.

Another recent branch of the literature investigates the impact of central bank communication on household inflation expectations. Coibion et al. (2022) show that the Fed's communication can have a significant effect on the household's inflation expectations when the communication is directly presented to the households. However, the effect of the communication is significantly damped when the same information is presented in the form of newspaper articles. This view is in line with Gardt et al. (2022) who demonstrate that households mostly hear about the ECB's monetary policy through television and newspapers (online and printed) and rarely by using direct sources like the ECB website. Similarly, Lamla & Vinogradov (2019) find that FOMC announcements have no impact on consumers' inflation perception and expectation, but that FOMC announcements increase the probability of consumers to receive news about the FOMC announcements.

Several studies investigate how central bank communication is perceived by the media. Berger et al. (2011) and Picault et al. (2022) show that the media's assessment of ECB policy decision is highly responsive to the content of ECB press conferences.

Similar to Picault et al. (2022), I adopt a framework proposed by Hayo & Neuenkirch (2015) where the media acts as a channel between central bank communication and the perception of monetary policy by financial markets. Following Nimark & Pitschner (2019) who formalize that agents delegate their information choice to the media rather than monitor all relevant events themselves. Hence, news act as a channel through which households receive information about the central bank communication. Building on these two frameworks, I assume that the media coverage of central bank communication acts as a channel between the central bank and households. I assume that households consume the news from the media about the central bank communication rather than directly from the central bank communication itself.

I am examining the impact of media coverage in newspapers on households' inflation expectation accuracy, specifically focusing on reporting of ECB press conferences and news about inflation. Firstly, I explore whether the inflation information presented in newspapers deviates from the inflation information in the press conferences by the ECB. Secondly, I investigate the extent to which the deviation of inflation information in media coverage and ECB press conferences contributes to explaining the errors in households' inflation expectations. This paper is, to my knowledge, the first to explicitly investigate the link between central bank communication and news coverage concerning household inflation expectations. Furthermore, it contributes to the literature by demonstrating how news about inflation deviates from the corresponding ECB communication, allowing for a better understanding of how central bank communication is conveyed to the general public and explaining the channel through which the general public reacts to central bank communication.

I built on the Bayesian learning model (Lamla & Lein, 2014) to examine the impact of the difference between inflation related information in media reporting and ECB press conferences on household's inflation expectations accuracy. To quantify the inflation related information in the ECB press conferences and news, I apply the lexicon driven procedure by Picault & Renault (2017) to create lexicons for inflation related information for the news and for the ECB press conferences. To measure how closely the media follows ECB press conferences in their news, I apply a similar procedure as Picault et al. (2022) and use dependency parsing to identify the grammatical structure of a text and filter out the share of news that reproduce ECB press conferences from general news about inflation. This allows me to measure how closely the media reporting follows the ECB's press conferences.

I find that the inflation information in both news and ECB press conferences shares a strong correlation with the current inflation rate in Germany. Moreover, the discrepancy in inflation reporting between news and ECB press conferences exhibits a significant correlation with the gap in inflation expectations between German households and professional inflation forecasts.

2. Model

I extend the Bayesian learning model for inflation expectation by Lamla & Lein (2014) by explicitly adding a channel for central bank communication to the media signal. I assume that the central bank sends a noisy signal for the future inflation π_{t+1} of the Euro area. Let c_t denote the signal that captures information about the central bank's inflation forecast that the central bank communicates to the media; that is, inflation-related information in press conferences. I further assume that the signal is normally distributed, $c_t \sim N(\pi_{t+1}, \sigma_t^c)$, with variance σ_t^c. This assumption implies that the central bank only sends one unambiguous signal at each period t rather than multiple conflicting signals at the same period,i.e., conflicting information in press conferences. In the absence of the central bank's communication, the media sends a noisy "baseline" signal about the rational forecast of inflation for future inflation π_{t+1} to households based on a number of media reports V, with $s^b_{\nu,t} \sim N(\pi_{t+1}, \sigma^{sb}_t)$ and a media bias α_t. I assume that α_t follows an AR(1) process with

$$\alpha_t = a_\alpha + \beta_\alpha * \alpha_{t-1} + \epsilon_\alpha \quad \epsilon_\alpha \sim N(0, \sigma_\alpha) \tag{1}$$

The media can decide how closely to follow the central bank's communication. Therefore, the media's inflation signal can be defined as:

$$s_{\nu,t} = (1 - \lambda_t)(s^b_{\nu,t} + \alpha_t) + \lambda_t c_t \quad 0 \leq \lambda_t \leq 1 \tag{2}$$

where λ_t denotes the weight that the media report places on the central bank's signal c_t relative to the "baseline" media signal $s^b_{\nu,t}$. A value of $\lambda_t = 0$ indicates that the media report completely ignores the central bank's communication, while $\lambda_t = 1$ describes a situation where the media only reproduces information from the central bank's communication. Given that $s_{\nu,t}$ is a linear combination of normal random variables and assuming that σ_α, σ^c_t and $\sigma^{sb}_{\nu,t}$ are independent, the media sends a normally distributed signal about π_{t+1} with $s_{\nu,t} \sim N(\mu^s_{\nu,t}, \sigma^s_{\nu,t})$, where:

$$\mu^s_{\nu,t} = (1 - \lambda_{\nu,t})(E(\alpha_t) + \pi_{t+1}) + \lambda_{\nu,t}\pi_{t+1} \tag{3}$$

$$\sigma^s_{\nu,t} = (1 - \lambda_{\nu,t})^2 \sigma^{s^b}_{\nu,t} + (1 - \lambda_{\nu,t})^2 \sigma^2_\alpha + \lambda^2_{\nu,t}\sigma^2_c \tag{4}$$

The households hold a prior belief $\gamma_t \sim N(\pi_t, \sigma^h_t)$ about the future inflation π_{t+1}. The households then update their beliefs based on the media's signal. Following the Bayesian updating rule, the households' posterior beliefs about future inflation π_{t+1} are:

$$k_i(\pi_{t+1}|s_{\nu,t}) \propto \Pi^V_{\nu=1} f(s_{\nu,t}|\pi_t)h(\pi_t) \tag{5}$$

where $k(.)$ is the posterior density given the medias signal, $f(.)$ is the conditional density of the observed media signal given the prior belief about π_t and $h(.)$ is the prior density of π_t. Assuming normality the posterior is normal distributed with $\pi^h_{t+1} \sim N(\mu^h_{t+1}, \sigma^h_{t+1})$ where μ^h_{t+1} and σ^h_{t+1} are the updated mean and variance of the household's belief about future inflation, respectively. To compute the updated mean and variance, the households combine their prior beliefs with the media's signal:

$$\mu_t = \rho_t \pi_t + (1 - \rho_t)\bar{\mu^s_t} \tag{6}$$

where

$$\bar{\mu^s_t} = V^{-1} \sum_{\nu=1}^{V} \mu^s_{\nu,t} = V^{-1} \sum_{\nu=1}^{V} (1 - \lambda_{\nu,t})(\alpha_t + \pi_{t+1}) + V^{-1} \sum_{\nu=1}^{V} \lambda_{\nu,t} \pi_{t+1} \tag{7}$$

$$= \pi_{t+1} + \alpha_t - \bar{\lambda}_t \alpha_t - \bar{\lambda}_t \pi_{t+1} + \bar{\lambda}_t \pi_{t+1} \tag{8}$$

$$= \alpha_t (1 - \bar{\lambda}_t) + \pi_{t+1} \tag{9}$$

and

$$\rho_t = \frac{\frac{1}{V}\sigma^s_t}{\sigma^h_t + \frac{1}{V}\sigma^s_t} \tag{10}$$

For $\bar{\lambda}_t = 0$, the mean would be $\mu_t = \rho_t \pi_t + (1 - \rho_t)(\alpha_t + \pi_{t+1})$. Hence, if the media completely ignores the central bank communication, the mean of the posterior is the weighted average of the prior and the biased media signal[1]. If the media perfectly reproduce the central bank communication, i.e., $\bar{\lambda}_t = 1$, the mean would be $\mu_t = \rho_t \pi_t + (1 - \rho_t)\pi_{t+1}$.

B_t is the difference between the professional inflation forecast and the household inflation expectations and $|B_t|$ is the absolute difference.

$$B_t = \mu_t - \pi_{t+1} = \rho_t \pi_t + (1 - \rho_t)(\alpha_t (1 - \bar{\lambda}_t) + \pi_{t+1}) - \pi_{t+1} \tag{11}$$

$$= (1 - \rho_t)(1 - \bar{\lambda}_t)\alpha_t + \rho(\pi_t - \pi_{t+1}) \tag{12}$$

$$|B_t| = |(1 - \rho_t)(1 - \bar{\lambda}_t)\alpha_t + \rho(\pi_t - \pi_{t+1})| \tag{13}$$

The effect of $\bar{\lambda}_t$ on B_t is

$$\frac{\partial B_t}{\partial \bar{\lambda}_t} = -(1 - \rho_t)\alpha_t \tag{14}$$

[1] This case is equivalent to the model from Lamla & Lein (2014) with media bias.

Hence, the effect of $\bar{\lambda}$ depends on the sign and size of the media bias α_t. Assuming a positive media bias like Lamla & Lein (2014) would mean that the effect from $\bar{\lambda_t}$ is negative.

If π_t is larger or close to π_{t+1}

$$(1-\rho_t)(1-\bar{\lambda}_t)\alpha_t + \rho(\pi_t - \pi_{t+1}) > 0 \tag{15}$$

and therefore

$$\frac{\partial |B|_t}{\partial \bar{\lambda_t}} = -(1-\rho_t)\alpha_t \tag{16}$$

Hence, an increase in λ_t will lead to a decrease in the inflation expectation error of the households. The intuition behind this is that if the prior inflation π_t is considerably larger or close to the rational forecast of π_{t+1}, households may overshoot their expectations based on the media bias. Consequently, increasing the weight given to the central bank communication helps to adjust the expectations downward. The opposite holds if π_{t+1} is significantly larger that π_t where α_t would act as a "correction" to increase the inflation expectation. From this follows my first two hypothesis

Hypothesis 1(a): *If the media's signal is affected by a media bias, an increase in the weight given to the central bank communication by the media has an positive effect on the household inflation forecasting expectations, depending on the size of the media bias.*

Hypothesis 1(b): *If the media's signal is affected by a media bias, an increase in the weight given to the central bank communication by the media has an ambiguous effect on the household inflation forecasting accuracy, depending on the size of the difference between the current inflation and the rational inflation forecast.*

Similarly, the media bias can affect the inflation forecasting accuracy of households. The partial effect of the media bias is

$$\frac{\partial B_t}{\partial \alpha_t} = (1 - \rho_t)(1 - \bar{\lambda_t}) > 0 \tag{17}$$

Hence, an increase of α_t will always increase the inflation expectations. For π_t larger or close to π_{t+1}

$$\frac{\partial |B|_t}{\partial \alpha_t} = (1 - \rho_t)(1 - \bar{\lambda_t}) > 0 \tag{18}$$

Hence, the media bias increase introduces an upward bias in the inflation expectations. In the case of $\pi_t << \pi_{t+1}$ media bias acts as a "correction" to counter the low prior inflation expectations if the actual future inflation is significantly higher. Hence, my next two hypothesis follow as

Hypothesis 2(a): *An increase of the media bias increases the inflation expectation of households. The size of the effect depends on weight given to the central bank communication*

Hypothesis 2(b): *An increase of the media bias by the media has an ambiguous effect on the household inflation forecasting accuracy depending on the size of the difference between the current inflation and the rational inflation forecast.*

Placeholder for uncertainty. Influence of Variance changes on the Inflation Expectation Bias

3. Data

My dataset consists of two distinct components: textual data from ECB press conferences and newspaper articles, and quantitative data on inflation and inflation expectations. Section 3.1 describes the textual data, and Section 3.2 outlines the quantitative data.

3.1. Textual Data

I collected two separate datasets for textual analysis: one for ECB communications and another for newspaper articles. The ECB's primary mode of communication with the media is through their press conferences, which are conducted in English every six weeks following the ECB's Governing Council monetary policy decisions. I gathered ECB press conference transcripts from the official ECB website from 2000 to 2022 using a web scraper. I excluded the Q&A section and other less significant sections, such as greetings and acknowledgments, retaining only the introductory statements.

The media news data was provided by Dpa, the largest German news agency. This dataset consists of over 7 million newspaper articles in German published by Dpa between 1991 and 2018. I filtered out articles unrelated to the economy and those solely focused on financial news, such as stock movement reports or business updates. Furthermore, I removed purely numeric information like tables from the articles. For a detailed description of the cleaning process, see [Cite Mariia, Philip, Kai, and me].

To isolate data relevant to inflation, I included only sentences containing the word "Inflation" and its synonyms, such as "Preisteigerung" or "Preiserhöhung" (price increase) [2].

To reduce the dimensionality of the remaining data, I applied several preprocessing steps commonly used in the literature to both textual datasets: lowercasing, removing punctuation, eliminating stopwords (e.g., and, but, or) that convey no information, and removing numbers.

[2]The words are: Inflation (Inflation), Inflationsrate (Inflation rate), Verbraucherpreisindex (Consumer Price Index, CPI), Lebenshaltungskosten (Cost of living), Geldentwertung (Currency devaluation), Teuerung (Price inflation), Preisanstieg (Price increase), Preiserhöhung (Price hike), Verteuerung (Increase in prices), Deflation (Deflation), Hyperinflation (Hyperinflation), Geldwert (Money value), Preisindex (Price index), Preisniveau (Price level), Kaufkraft (Purchasing power), and Warenkorb (Basket of goods).

3.2. *Quantitative Data*

I use the quarterly year-on-year HICP growth rate for Germany, sourced from the Eurostat database, as a measure of inflation.
For the rational inflation forecast, I employ the ECB surveys of professional forecasters. Each quarter, the ECB surveys professional forecasters about their year-on-year inflation expectations. To obtain the final forecast, I calculate the average of all forecasts. The survey data covers the period from the first quarter of 1999 to the last quarter of 2022.
For household inflation expectations, I utilize the monthly European Commission's Business and Consumer Survey. In this survey, German households are asked about their expectations regarding price changes in the next 12 months. They can anticipate faster rising prices, prices rising at a consistent pace, prices rising more slowly, prices remaining unchanged, or prices decreasing.
To transform qualitative survey responses into quantitative inflation expectations, I apply a rolling-window regression-based approach by Lahiri & Zhao (2015). This approach is based on an extended version of Carlson & Parkin (1975) by Berk (1999) and has been simplified by Rosenblatt-Wisch & Scheufele (2015). A detailed explanation can be found in Section Appendix A.
Although quantifying qualitative survey data is not as optimal as using quantitative survey results, it is necessary due to the lack of publicly available quantitative survey-based inflation expectations data for German households.

4. Text Classification

In this section, I explain my methodology for quantifying ECB communication and newspaper articles. I divide my dataset into individual sentences and classify each sentence according to predefined categories. ECB press conferences fall into three categories, each consisting of three classes. The first two categories, adapted from Picault & Renault (2017), characterize the monetary stance as monetary hawkish, monetary neutral, or monetary dovish, and represent the economic outlook as positive, neutral, or negative. I introduce a third category that conveys the inflation outlook as increasing, steady, or decreasing, enabling a direct comparison between inflation expectations communicated in ECB press conferences and those expressed in the news media.

News articles are classified into two categories, each with three classes. The first category describes the economic impact of inflation reported in the news as positive, neutral, or negative, while the second category indicates the direction of inflation as increasing, neutral, or decreasing.

To achieve the text classification two main approaches are used in the literature: Machine Learning text classification and Lexicon-based text classification. Machine Learning (ML) text classification utilizes complex probabilistic models, usually trained on large datasets. Modern ML models are able to take into account the grammatical and syntactical structure of texts. However, the complexity of these models comes with a drawback. Modern deep-learning models like BERT (Devlin et al 2018) are "black-boxes" and require large training datasets.

Lexicon-based text classification utilizes a lexicon to classify the category of a given text. A lexicon is a collection of words or phrases associated which specific category classes. This approach relies on the frequency, or weights of the words or phrases in the lexicon to determine the category of a given text.

The simple implementation and transparency of Lexicon-based sentiment classification has made it widely used in the economic and finance literature, e.g., Shapiro et al. (2022), Loughran & McDonald (2011), Barbaglia et al. (2022), Nyman et al. (2021), Ardia et al. (2019). The most prominent lexicon for sentiment classification of texts from the finance and economic domain was created by Loughran & McDonald (2011)). Picault & Renault (2017) developed a lexicon for classifying ECB press conferences and measuring the conveyed economic outlook and monetary stance . However, these lexicons were not created for classifying inflation directions and sentiment with regard to inflation. Hence, I implement a similar approach like Shapiro et al. (2022) and Picault & Renault (2017) and create data-driven lexicons based on the training datasets created by me. The lexicon is then used to classify each sentence in the ECB press conferences according to the three mentioned categories and each of the newspaper sentences according to the two mentioned categories.

I created two training datasets: one for the newspaper articles and one for the ECB press conferences. Due to the different categories and different styles of these two types of text, I created two training datassets.

I manually classified 3000 randomly drawn sentences from the ECB press conferences according to the three categories, with each sentence being labeled with one of the three classes for each category. Similarly, 3000 sentences were randomly drawn from the news corpus and labeled based on the two categories. I then used these datasets to create one lexicon for the ECB press conferences and for the newspaper articles.

The lexicons are created by calculating the degree to which each word in the training dataset is associated with a class. The association is calculated using pointwise mutual information (PMI). Following Church and Hanks (1990) PMI is defined as

$$PMI(c,w) = log_2(\frac{p(c,w)}{p(c)p(w)}) \tag{19}$$

where $p(w)$ is the words share of total words in the training dataset, $p(c)$ is the total share of sentences belonging to class c and $p(c,w)$ is the words share in class c of total words. The PMI measure how closely a is associated with a class. This formula is applied to all words for all categories in the two training datasets.

The lexicons are then used to to assign a category specific score to each press conference and newspaper sentences. Each ECB press conference is treated as a independent text and all newspapers sentences from one month are concatenated to one text for each month. Then I calculate a score for each category and score by.

$$score_i^c = \frac{\sum_{j=1}^{n}(PMI(c_3,w_{i,j}) - PMI(c_1,w_{i,j}))}{\sum_{j=1}^{n}(PMI(c_3,w_{i,j}) + PMI(c_2,w_{i,j}) + PMI(c_1,w_{i,j}))} \tag{20}$$

where $score_i^c$ is the score of text i specific to category c_g. g denotes the class in the respective category.For the ECB monetary categories these are c_3 for hawkish, positive economic outlook and increasing inflation. c_1 stands for dovish, negative economic outlook and decreasing inflation. c_2 denotes the neutral class. For the newspaper data c_3 stand for increasing inflation and positive sentiment. c_1 for decreasing inflation and negative sentiment. $PMI(c_3, w_{i,j})$, $PMI(c_2, w_{i,j})$, and $PMI(c_1, w_{i,j})$ are the PMI values of word $w_{i,j}$ in text i of the respective class, and n is the number of words in the text.

To improve the text classification, I follow Shapiro et al. (2022) and implement negations. Negation words like "not" directly impact the class of a sentence, e.g., "the inflation is rising" vs. "the inflation is not rising". I follow CITE and handle these negations by multiplying the PMI-value of each word by -1 if the word is within a three word window of a negation word.

For the newspaper articles, I have therefore two monthly indizes. One for the direction of the inflation and one for the sentiment regarding the inflation. I combine both these indizes to one by multiplying them with each other. I call the resulting index Inflation News Index (INI).

The (ECB) press conferences, occur eight times a year. To convert the press conference-based indizes to a monthly frequency, I create a monthly time series covering the entire period for which I have press conference data. For each month, I assign the index value from the most recent press conference preceding that month. This approach assumes that the index remains constant between press conferences and only changes when a new press conference is released. By using this method, I obtain a monthly index that reflects the information available to market participants at each point in time.

4.1. Measuring how closely the Media follows Central Bank Communication

Dependency parsing is a natural processing techniques that identifies the grammatical structure of a given text.

I apply the Stanford stanza dependency in python to all sentences in which either the ECB or one of the past or present ecb governing councile members are mentioned. I distinct between sentences where the ECB is directly or indirectly quoted from sentences in which the journalist makes a judgment about the ECB's actions. I only keep the first and discard the later. To achieve this

I filter out the sentence with two rules. First, I keep all sentences in which the ECB or one of its governing council members is the main subject. Second, I keep all sentences in which the ECB or one of its governing council is the object and a verb which corresponds to communication like sagen (saysing) or berichten (reporting)[3].

This approach is close to Picault et al. (2022), expect that they use part-of-speech tagging instead of dependency parsing to identify the corresponding sentences.

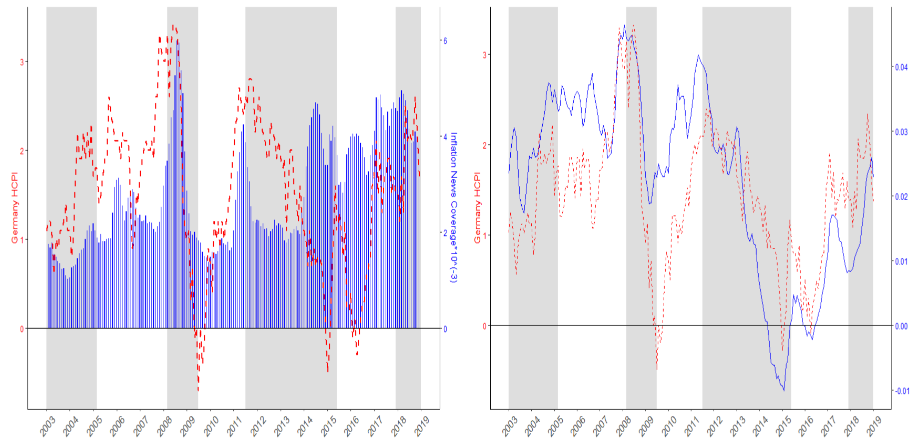
5. Descriptive Analysis

I relate my news index and the amount of media coverage to inflation. Prior literature has shown that media coverage can have a significant impact on inflation expectations and consumer behavior Carroll (2003). In general, inflation-related news strongly correlates with inflation itself. As illustrated in Figure A.4(a), high inflation phases in 2008 and 2011/2012 coincide with short peaks in inflation-related news, while periods of lower inflation in between have less media coverage.

However, after 2013, media coverage peaked during the low inflation phase from 2014 to 2017 and remained elevated afterward. This is noteworthy as it contrasts with the fiding of positive correlation between media coverage and inflation from Lamla & Lein (2014). Figure A.4(b) demonstrates that the INI correlates with inflation, with the notable exception of the inflation decrease during the Great Recession, which saw only a moderate decline in the INI. In contrast, the low inflation phase from 2014 to 2016 is accompanied by a dip in the INI lower than during the Great Recession.

Figure 2 illustrates the strong correlation between Eurozone inflation, the ECB inflation index, and the ECB staff inflation projection. In 2006, the ECB Inflation Index reached a peak near the height of the Great Recession, even

[3]These words are: sagen (say), sehen (see), gehen (go), berichten (report), meinen (mean), erwarten (expect), vorhersagen (predict), rechnen (calculate)



(a) Media Coverage and Inflation

(b) INI and Inflation

Figure 1: Media Reporting and Inflation: Figure (a) depicts the monthly year-on-year inflation for Germany in red and the media coverage in blue. Media coverage is defined as the number of inflation-related sentences divided by the total number of sentences to account for the varying number of articles in each year. Figure (b) depicts the monthly year-on-year inflation for Germany in red and the INI in blue.

though inflation during this period was relatively lower. Similarly to the INI, the low inflation experienced at the beginning of 2009 was accompanied by a considerably smaller dip in the ECB inflation index compared to the 2014 dip. In this case, the ECB inflation index led Eurozone inflation from 2011 to 2014.

Figure 3 shows that media bias correlates with the media bias. Especially after the Great Recession the expectation gap and the media bias dropped and remained low until they increased again after 2014.

6. Econometric Framework

To test my two hypothesis, I examine the impact of the media bias and media-central bank alignment on the inflation expectation error. To do this, I estimate the following equation

$B_t = \beta_0 + \beta_1 \tilde{V}_t + \beta_2 \tilde{\alpha}_t + \beta_3 \tilde{\lambda}_t + \beta_3 \pi_{t-1} + \beta_4 B_{t-1} + \epsilon_t$

where β_0 is a constant, $\tilde{V}_t$ is the share of inflation related sentences in the

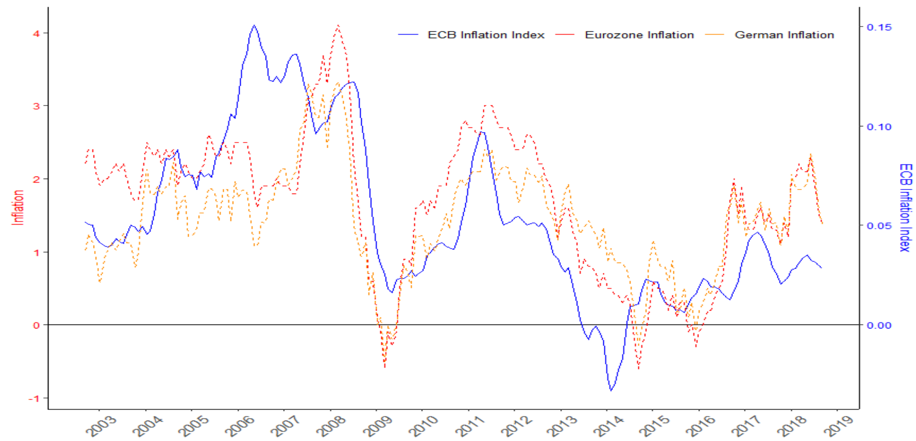


Figure 2: The plot depicts the monthly year-on-year inflation in the Eurozone in red, alongside the ECB inflation index in blue.

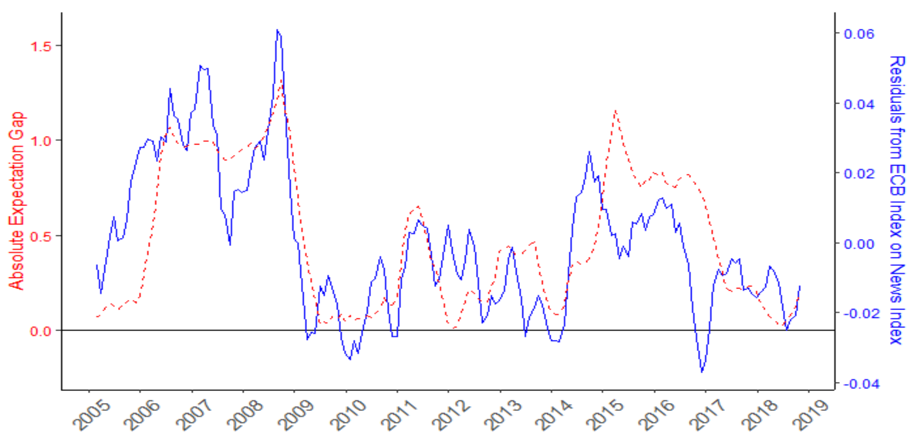


Figure 3: The figure shows the absolute expectation gap between proffesional inflation forecasts and quantified household inflation expectations in red and the regression residuals from the ECB inflation index on the INI in blue.

media divided by the number of all sentences, $\tilde{\alpha_t}$ is the residual from the linear regression of the ECB inflation index on the INI and media coverage, $\tilde{\lambda_t}$ is the number of sentences in which the ECB is directly or indirectly cited divided by the number of all sentences. The equation is estimated via OLS using Newey-West standard errors.

7. Results

Table 1: ECB - Stm

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		0.032	-0.019	-0.035	-0.077
		(0.112)	(0.109)	(0.105)	(0.107)
π_{t-1}		-0.005	-0.016	-0.023	-0.032
		(0.021)	(0.022)	(0.019)	(0.021)
$\tilde{V}_t$	-7.290	-7.613	-16.667	5.391	-4.107
	(15.445)	(15.842)	(16.480)	(13.313)	(13.609)
$\tilde{\alpha}_t$			1.638*		1.521*
			(0.873)		(0.824)
$\tilde{\lambda}_t$				-3.624***	-3.321***
				(1.179)	(1.030)
$Constant$	0.254***	0.253***	0.281***	0.351***	0.369***
	(0.033)	(0.051)	(0.057)	(0.053)	(0.058)
R^2	0.004	0.005	0.101	0.088	0.170

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2: ECB - Berk 1

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		0.196**	0.187**	0.190**	0.183**
		(0.088)	(0.086)	(0.092)	(0.091)
π_{t-1}		-0.286***	-0.290***	-0.289***	-0.293***
		(0.040)	(0.039)	(0.040)	(0.041)
$\tilde{V}_t$	-57.757	-75.719***	-78.287***	-73.604***	-76.451***
	(42.787)	(25.444)	(25.221)	(25.929)	(25.463)
$\tilde{\alpha}_t$			0.507		0.493
			(1.256)		(1.255)
$\tilde{\lambda}_t$				-0.540	-0.451
				(1.404)	(1.396)
Constant	0.540***	0.837***	0.848***	0.854***	0.861***
	(0.115)	(0.093)	(0.093)	(0.110)	(0.113)
R^2	0.053	0.552	0.554	0.553	0.555

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

8. Conclusions

References

Ardia, D., Bluteau, K., & Boudt, K. (2019). Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting*, *35*, 1370–1386.

Barbaglia, L., Consoli, S., & Manzan, S. (2022). Forecasting with economic news. *Journal of Business & Economic Statistics*, (pp. 1–12).

Batchelor, R., & Orr, A. B. (1988). Inflation expectations revisited. *Economica*, *55*, 317–331.

Berger, H., Ehrmann, M., & Fratzscher, M. (2011). Monetary policy in the media. *Journal of Money, Credit and Banking*, *43*, 689–709.

Table 3: ECB - Berk 5

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		0.463***	0.409***	0.371***	0.324***
		(0.107)	(0.102)	(0.102)	(0.096)
π_{t-1}		0.072	0.061*	0.056	0.047
		(0.049)	(0.036)	(0.045)	(0.033)
$\tilde{V}_t$	-56.125**	-25.989	-41.558***	-14.878	-30.443**
	(25.206)	(17.162)	(14.396)	(14.779)	(12.348)
$\tilde{\alpha}_t$			2.203**		2.113***
			(0.865)		(0.799)
$\tilde{\lambda}_t$				-4.666***	-4.401***
				(0.964)	(0.820)
$Constant$	0.455***	0.137**	0.177***	0.273***	0.304***
	(0.064)	(0.055)	(0.055)	(0.050)	(0.050)
R^2	0.089	0.397	0.468	0.448	0.513

*** p < 0.01, ** p < 0.05, * p < 0.1.

Berk, J. M. (1999). Measuring inflation expectations: a survey data approach. *Applied Economics*, *31.11*, 1467–1480.

Blinder, A., Ehrmann, M., De Haan, J., & Jansen, D.-J. (2017). Necessity as the mother of invention: Monetary policy after the crisis. *Economic Policy*, *32.92*, 707–755.

Blinder, A., Ehrmann, M., De Haan, J., & Jansen, D.-J. (2022). Central bank communication with the general public: Promise or false hope? *Journal of Economic Literature (forthcoming)*, .

Carlson, J. A., & Parkin, M. (1975). Inflation expectations. *Economica*, *42.166*, 123–138.

Table 4: ECB - Quant

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		0.362	0.375	0.289	0.305
		(0.641)	(0.539)	(0.545)	(0.447)
π_{t-1}		0.247	0.174	0.189	0.129
		(0.394)	(0.346)	(0.341)	(0.308)
$\tilde{V}_t$	121.867	312.588	270.628	375.381	335.130
	(552.031)	(290.579)	(268.372)	(281.803)	(273.899)
$\tilde{\alpha}_t$			7.704*		6.691*
			(4.363)		(3.644)
$\tilde{\lambda}_t$				-26.143**	-24.557**
				(10.258)	(10.015)
Constant	2.550***	1.130	1.200	1.817*	1.837**
	(0.644)	(1.110)	(1.018)	(0.982)	(0.843)
R^2	0.019	0.266	0.310	0.355	0.388

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of Economics*, *118.1*, 269–298.

Coibion, O., Gorodnichenko, Y., & Weber, M. (2022). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, *130*, 1537–1584.

Van der Cruijsen, C., Jansen, D.-J., & de Haan, J. (2015). How much does the public know about the ecb's monetary policy? evidence from a survey of dutch households. *International Journal of Central Banking*, *11*, 169–218.

Ehrmann, M., Pfajfar, D., & Santoro, E. (2017). Consumers' attitudes and their inflation expectations. *International Journal of Central Banking*, *13.1*, 225–259.

Gardt, M., Siria, A., Mee, S., & Glöckler, G. (2022). Ecb communication with the wider public. *Economic Bulletin Articles*, *8*.

Hayo, B., & Neuenkirch, E. (2015). Self-monitoring or reliance on media reporting: How do financial market participants process central bank news? *Journal of Banking & Finance*, *59*, 27–37.

Lahiri, K., & Zhao, Y. (2015). Quantifying survey expectations: A critical review and generalization of the carlson–parkin method. *International Journal of Forecasting*, *31.1*, 51–62.

Lamla, M. J., & Lein, S. M. (2014). The role of media for consumers' inflation expectation formation. *Journal of Economic Behavior & Organization*, *106*, 62–77.

Lamla, M. J., & Vinogradov, D. V. (2019). Central bank announcements: Big news for little people? *Journal of Monetary Economics*, *108*, 21–38.

Larsen, V. H., Thorsrud, L. A., & Zhulanova, J. (2021). News-driven inflation expectations and information rigidities. *JJournal of Monetary Economics*, *117*, 507–520.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of finance*, *66*, 35–65.

Nimark, K. P., & Pitschner, S. (2019). News media and delegated information choice. *Journal of Economic Theory*, *181*, 160–196.

Nyman, R., Kapadia, S., & Tuckett, D. (2021). News and narratives in financial systems: exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control*, *127*.

Picault, M., Pinter, J., & Renault, T. (2022). Media sentiment on monetary policy: Determinants and relevance for inflation expectations. *Journal of International Money and Finance*, *124*.

Picault, M., & Renault, T. (2017). Words are not all created equal: A new measure of ecb communication. *Journal of International Money and Finance*, *79*, 136–156.

Rosenblatt-Wisch, R., & Scheufele, R. (2015). Quantification and characteristics of household inflation expectations in switzerland. *Applied Economics*, *47.26*, 2699–2716.

Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2022). Measuring news sentiment. *Journal of econometrics*, *228*, 221–243.

Appendix A. Quantification of Inflation Expectations

The European Commission's Business and Consumer Survey consumers are asked if they expect prices to fall, stay the same, increase slower than before, increase at the same rate or increase at a higher rate in the coming 12 months and in the last 12 months.

I use the rolling-window-regression approach by Lahiri & Zhao (2015) which is based on an extended version of the Carlson-Parkin method (Carlson & Parkin, 1975) by Berk (1999). The extended version of Berk (1999) does not impose unbiasedness of inflation expectations. Instead the current perceived inflation rate is directly linked to the expected inflation rate.

By assumption each responded i forms a subjective probability distribution for individuals percentage price changes y_{it} over the next twelve months. Let $f_i(y_it)$ be the subjective probability distribution with mean μ_{it} and variance σ_{it}. Following Batchelor & Orr (1988) method assumes for a survey with five possible answers that the respondent answers prices in the future increase, increase at the same rate, increase at a slower rate, stay the same or fall according as $y_{it} < -\delta_{it}^L, -\delta_{it}^L < y_{it} < \delta_{it}^U, \delta_{it}^L < y_{it} < \delta_{it}^U, \delta_{it}^U < y_{it} < \lambda_{it}, \lambda_{it} < y_{it}$ Nochmal überprüfen, Ist Notation von Batchelor & Orr (1988) besser?. Based on the survey responses the corresponding aggregate probabilities can be formulated as $P(y < -\delta_{it}^L) = A_t, P(y < \delta_{it}^U) - P(y > -\delta_{it}^L) = B_t, P(y < \delta_{it}^U) - P(y > \delta_{it}^L) = C_t, P(y < \lambda_{it}) - P(y > \delta_{it}^U) = D_t$. I denote a_t, b_t, c_t and d_t the abscissae of the standard logistical distribution function corresponding to the cumulative probabilities $A_t,\ A_t + B_t,\ A_t + B_t + C_t$ and $A_t + B_t + C_t + D_t$. The mean expected inflation rate $\mu_t = E_t\pi_{t+12}$ can then be formulated as

$$\mu_t = \lambda_t \frac{(a_t + b_t)}{(a_t + b_t - c_t - d_t)}$$

Similarly, I denote a_t', b_t', c_t' and d_t' as the abscissae of the standard logistical distribution function for the perceived inflation. Assuming that the response threshold λ_{it} Nochmal überprüfen is the same for expected and perceived inflation, the perceived inflation can be formulated as

$$\mu'_t = \lambda_t \frac{(a'_t + b'_t)}{(a'_t + b'_t - c'_t - d'_t)}$$

For the choice of the scaling parameter λ_t I use the rolling window based regression by Lahiri & Zhao (2015). Following Rosenblatt-Wisch & Scheufele (2015) running the regression

$$\pi_t = \lambda \frac{(a'_t + b'_t)}{(a'_t + b'_t - c'_t - d'_t)} + u_t$$

using a sample window of $t - w + 1tot$ implies

$$\hat{\lambda}_t = \frac{\sum_{k=t-w+1}^{t} (a'_t + b'_t)/(a'_t + b'_t - c'_t - d'_t)\pi_t}{\sum_{k=t-w+1}^{t} (a'_t + b'_t)/((a'_t + b'_t - c'_t - d'_t))^2}$$

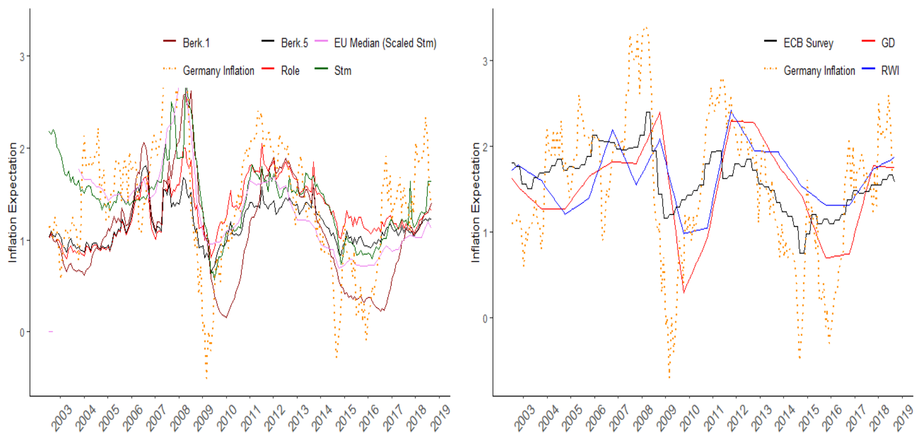
where w is the size of the rolling window. Lahiri & Zhao (2015) choose w as nine years for the inflation expectations of the Michigan Consumer Survey. I follow Lahiri & Zhao (2015) and also choose nine years for w. The resulting household inflation expectations and the professional inflation expectations are shown in figure **??**.

Robustness Check for different windows?

TO CLOSE TO Rosenblatt-Wisch and Scheufele 2014??

Appendix B. Alternative Survey Quantification Methods and Professional Forecasts

Appendix C. Additional Rules for Lexicon Text Classification



(a) Household Inflation Expectations

(b) Professional Inflation Expectations

Figure A.4: Inflation Expectations.

Table B.5: Inflation Expectation Gap: RWI - Stm

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		0.052	0.057	0.041	0.047
		(0.126)	(0.125)	(0.129)	(0.128)
π_{t-1}		0.103**	0.102**	0.108***	0.107***
		(0.043)	(0.043)	(0.040)	(0.041)
$\tilde{V}_t$	41.134*	48.032***	47.606***	43.221**	42.613**
	(22.475)	(18.037)	(18.098)	(18.726)	(18.753)
$\tilde{\alpha}_t$			0.111		0.138
			(0.721)		(0.722)
$\tilde{\lambda}_t$				1.136	1.155
				(1.355)	(1.344)
$Constant$	0.340***	0.176**	0.176**	0.156**	0.155**
	(0.052)	(0.081)	(0.082)	(0.077)	(0.078)
R^2	0.067	0.190	0.190	0.194	0.195

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.6: Inflation Expectation Gap: RWI - Berk 1

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		0.155	0.131	0.140	0.123
		(0.283)	(0.257)	(0.234)	(0.211)
π_{t-1}		-0.140**	-0.139**	-0.121*	-0.121*
		(0.068)	(0.066)	(0.065)	(0.063)
$\tilde{V}_t$	25.468	5.626	9.958	-14.568	-10.634
	(32.896)	(40.084)	(38.428)	(39.098)	(37.603)
$\tilde{\alpha}_t$			-0.820		-0.635
			(1.516)		(1.508)
$\tilde{\lambda}_t$				4.931***	4.791***
				(1.634)	(1.584)
$Constant$	0.592***	0.702***	0.714***	0.609***	0.621***
	(0.119)	(0.229)	(0.207)	(0.222)	(0.201)
R^2	0.013	0.199	0.206	0.241	0.245

*** p < 0.01, ** p < 0.05, * p < 0.1.

Appendix D. Open Problems

Table B.7: Inflation Expectation Gap: RWI - Berk 5

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		-0.132	-0.152	-0.139	-0.156
		(0.192)	(0.186)	(0.186)	(0.185)
π_{t-1}		0.158***	0.162***	0.167***	0.170***
		(0.045)	(0.042)	(0.042)	(0.041)
$\tilde{V}_t$	-39.960	-31.620	-29.295	-40.853	-38.112
	(33.795)	(25.727)	(24.489)	(26.275)	(24.478)
$\tilde{\alpha}_t$			-0.746		-0.669
			(1.338)		(1.340)
$\tilde{\lambda}_t$				2.133	1.982
				(1.632)	(1.550)
Constant	0.485***	0.323**	0.328**	0.282**	0.289**
	(0.086)	(0.135)	(0.135)	(0.138)	(0.138)
R^2	0.044	0.296	0.304	0.307	0.314

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Assuming that agents can't observe α and λ_t if they both follow an AR(1) process. Would I need to condition the likelihood on them and then also assume priors for α and λ_t or cold I just ignore their time varying nature for the formation of my posterior? It would imply that households do not form expectations about the formation of alpha and lambda and simply take the as given.

The steady state for

$$E(\pi_{t+1}) = (1 - \lambda) * (\alpha_t + \Psi_t) + \lambda * \Psi_t$$

is only defined for a α_t without a constant (steady state of zero), assuming that $\Psi_t = E(\pi_{t+1})$.

α_t can be exogenous and follow an AR(1) process. Hence, α_t would be following

Table B.8: Inflation Expectation Gap: GD - Stm

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		0.112	0.016	0.110	0.016
		(0.149)	(0.144)	(0.149)	(0.145)
π_{t-1}		0.071	0.083*	0.072	0.083*
		(0.054)	(0.045)	(0.054)	(0.045)
$\tilde{V}_t$	-25.775	-18.217	-14.263	-19.757	-14.424
	(17.555)	(16.962)	(15.559)	(18.211)	(16.574)
$\tilde{\alpha}_t$			-1.782***		-1.781***
			(0.567)		(0.571)
$\tilde{\lambda}_t$				0.353	0.036
				(1.397)	(1.440)
$Constant$	0.325***	0.188**	0.202***	0.181**	0.202***
	(0.047)	(0.079)	(0.075)	(0.080)	(0.077)
R^2	0.029	0.104	0.169	0.105	0.169

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

a normal distribution. However, would it follow the conditional (on previous α_t) or unconditional normal distribution?

If Lambda also follows an AR(1) process I have a problem with the normal distributions; Normal times normal. Could I just ignore that and treat lambda as a parameter and change it to an AR(1) process after I formulated the mean with the Bayesian learning framework?
Or come up with a conditional distribution and estimate mean numerically? Is that how you would do it for a DSGE model?

Assuming that V is not a constant but rather an exogenous variable which

Table B.9: Inflation Expectation Gap: GD - Berk 1

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		0.193	0.180	0.164	0.155
		(0.176)	(0.176)	(0.182)	(0.183)
π_{t-1}		0.040	0.043	0.052	0.053
		(0.038)	(0.038)	(0.032)	(0.033)
$\tilde{V}_t$	-2.105	-2.325	-0.318	-15.398	-13.481
	(18.141)	(19.385)	(18.836)	(19.883)	(19.435)
$\tilde{\alpha}_t$			-0.504		-0.393
			(1.030)		(0.995)
$\tilde{\lambda}_t$				3.171*	3.085*
				(1.644)	(1.639)
$Constant$	0.442***	0.301**	0.303**	0.249**	0.252**
	(0.053)	(0.120)	(0.125)	(0.113)	(0.115)
R^2	0.0002	0.037	0.044	0.081	0.085

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

follows an AR(1) process.

$$V_t = V_c onst + rho_V * V_{t-1} + \epsilon_V \tag{D.1}$$

Would it influence the posterior distribution? Would net log-linearization around steady state for inflation expectations.

Table B.10: GD - Berk 5

	(1)	(2)	(3)	(4)	(5)
B_{t-1}		0.032	-0.015	0.031	-0.015
		(0.126)	(0.122)	(0.128)	(0.122)
π_{t-1}		0.204***	0.215***	0.207***	0.216***
		(0.059)	(0.058)	(0.059)	(0.058)
$\tilde{V}_t$	-34.755	-21.298	-17.411	-24.066	-18.831
	(25.552)	(18.306)	(17.546)	(19.283)	(18.434)
$\tilde{\alpha}_t$			-1.420		-1.407
			(1.008)		(1.023)
$\tilde{\lambda}_t$				0.651	0.325
				(1.311)	(1.284)
$Constant$	0.470***	0.165**	0.173**	0.152**	0.166**
	(0.082)	(0.071)	(0.069)	(0.074)	(0.069)
R^2	0.034	0.381	0.409	0.382	0.409

*** p < 0.01, ** p < 0.05, * p < 0.1.