

Different Newspapers—Different Expectations*

—*Preliminary and Incomplete*—

Sarah Arndt
Heidelberg University

April 24, 2024

Abstract

I explore the extent to which the inflation signals emitted by different types of newspapers can explain household's inflation expectations in Germany. Constructing newspaper-specific indicators using the articles' text data and the large language model *gpt-3.5-turbo-1106*, I can find heterogeneity in the informativeness regarding the expectations of different socioeconomic groups, depending on the genre of the newspaper (i.e., tabloid vs. reputable newspapers). The tabloid's inflation expectation indicator performs better in predicting expectations of low income and lower education households, while the contrary holds true for the more serious newspapers. I then employ local projections to estimate the newspaper indices' responses to a monetary policy shock. This again reveals heterogeneity with the tabloid's sentiment displaying an immediate decrease while other newspaper's responses are smaller and lagged.

Keywords: Inflation expectations, text mining, forecasting, monetary policy, LLM, ChatGPT

JEL-Codes: E31, E58, C53, C55

*I thank the FAZ and SZ archives for allowing me to retrieve FAZ and SZ newspaper articles from the respective online database. Their cooperation is gratefully acknowledged.

1 Introduction

Households observe various inflation signals ranging from grocery prices to newspaper articles and social media posts. They use these sources of information to form their expectations about future inflation. Sometimes, these expectations become de-anchored and exceed what actual levels of inflation would suggest. This happened for example in Germany, in mid 2021. Even though inflation was below target in 2020, inflation expectations measured by the European Commission’s (EC) *Business and Consumer Survey* were already increasing and attained its highest levels since 2011 (see Figure 1). At the same time, Germany’s yellow press suddenly increased their reporting on inflation, as shown by the blue shaded area in Figure 2, representing the absolute weekly number of articles related to the topic of inflation published by Germany’s largest boulevard newspaper—*Die Bild*. This press is known for its polarizing style of reporting and published headlines such as “*The inflation monster is back!*” (May 8, 2021), even though inflation was barely exceeding 2% back then, and one month later: “*Inflation is eating away our savings*” (June 2, 2021), fueling households’ fear of inflation. There is the potential risk that the public opinion had been distorted by this sudden increase of dramatized media coverage.

This paper takes a closer look at the different signals towards inflation emitted by the yellow press, represented by Germany’s largest newspaper in-terms of sold units—*Die Bild*, and more reputable daily German papers like the *Frankfurter Allgemeine Zeitung* (FAZ) and *Die Süddeutsche Zeitung* (SZ), being the two most popular representatives. I collect inflation related articles of these three newspapers and analyze their sentiment towards inflation with the large language model gpt-3.5-turbo-1106 developed by OpenAI. Out of the raw sentiment scores per article, I use different methods to compute overall sentiment indices per newspaper.

As it is well established, that some socioeconomic groups of households have persistently different inflation expectations (Meeks and Monti, 2023), I test, if different newspaper sentiment indices can account for this heterogeneity. Therefore, I compare the newspaper indices with household survey inflation expectations differentiated by age, income, education and gender of the survey participants and find heterogeneity in contemporaneous correlations, suggesting that expectation extracted from the yellow press rather resonate with lower education households while the opposite holds true for FAZ and SZ.

In a second step, I analyze the forecasting power of the three news indices targeting the different household sub-categories’ survey expectations by conducting an out-of-sample forecasting exercise. Again, I can find heterogeneity among the indices suggesting that the Bild index is more suitable for predicting expectations of lower income and lower education survey participants, while the opposite holds true for the other two newspaper indices. The latter outperform the Bild index in predicting inflation in Germany. In general, all newspaper indices carry some informativeness regarding household survey expectations as

including them in a forecasting model reduces the root mean squared error in most cases. This makes them an attractive alternative measure of household inflation expectations due to its high-frequency real-time availability.

I investigate the effect of a euro zone monetary policy shock on the three newspaper inflation expectation indices. While FAZ and SZ display a more muted and similar response, the Bild index reacts with an immediate reduction in overall sentiment in the days after a shock happened. This observation suggests, that either monetary policy is more effective in taming the Bild’s negative sentiment towards inflation or could result from the newspaper’s higher volatility in reporting on inflation, peaking around the monetary policy meetings.

Lastly, I explore another possible source of heterogeneity by comparing the responses of household expectations resulting from a shock to the sentiment of the three newspapers. Expectations survey participants with higher income and a higher education level are most responsive to a shock in FAZ and SZ sentiment. In contrast, Bild sentiment shocks are mostly reaching households in the second income quartile, even though overall heterogeneity is less pronounced in the case of the Bild. This shows, that some of the heterogeneous informativeness of the indicators could be explained by their differing ability to affect household expectations.

So far, as Meeks and Monti (2023) highlight, the existing literature and policy discussions mainly focus on aggregate household inflation expectations. Previous studies eliciting inflation expectations indices from text data, specifically posts on Twitter (today known as *X*) (see, e.g., Angelico et al., 2022; Born et al., 2023) test the informativeness of their indicators only against aggregate household survey expectations. One contribution of this paper is to differentiate the survey expectations across various socioeconomic groups and to test the informativeness of different text sources for these specific households. With this I shed light on possible drivers of heterogeneity in household inflation expectations.

There is already a growing body of literature that employs newspaper text data and machine learning methods to extract sentiment and expectations, often used for forecasting macro variables—see for example Bybee (2023), Shapiro et al. (2022), Picault et al. (2022), or Ellingsen et al. (2022). Yet, these papers take single or few reputable newspapers as a representative for all news data and neglect that different genre of newspaper might emit different signals that reach people differently. In a similar vein to this paper, Menz and Poppitz (2013) explain socioeconomic differences in inflation expectations in Germany with heterogeneity in news consumption. In contrast, they rely on manually labeled (into positive and negative) newspaper articles including the Bild (but not FAZ and SZ) and television news on a sample spanning from 1999 to 2010. In contrast, their aim is not to extract inflation expectation sentiments from the news data, but to test the epidemiological model introduced by Carroll (2003) to see if tone and news coverage brings household expectations closer to those of professional forecasters.

The paper is structured in the following way: Section 2 describes the newspaper and survey data set. Section 3 outlines construction of the newspaper inflation expectations indices, followed by an analysis of their informativeness in Section 4. I present the effects of a monetary policy shock on the newspaper indices in Section 5, the responses of survey expectations after a shock in news sentiment in Section 6 and Section 7 concludes.

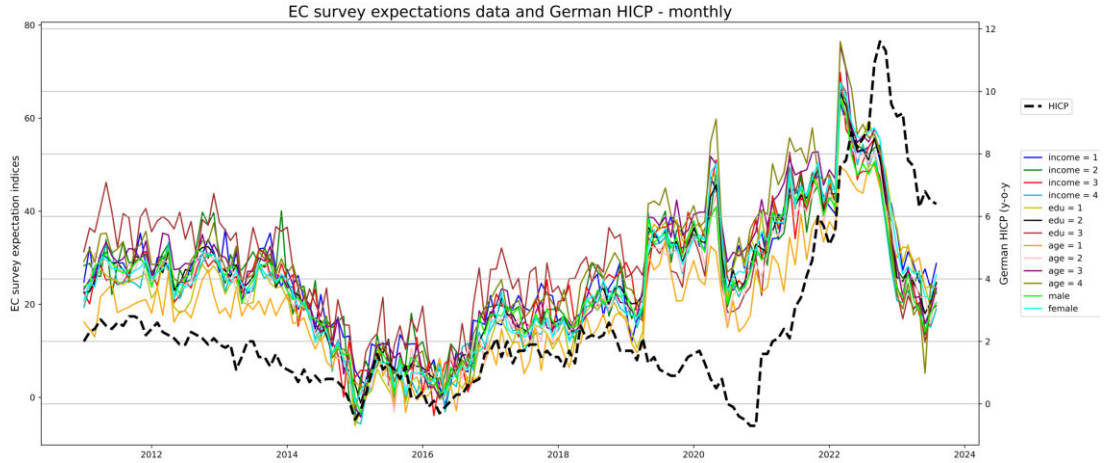


Figure 1: Left axis: EC survey expectation indices for different socioeconomic groups. Right axis: monthly year-on-year growth rates of German HICP (red line). Sample span: January 2011 to August 2023.

2 Data Sources

2.1 Newspaper Data

In Germany, there is only one daily popular nationwide yellow press newspaper—*Die Bild* (Bild) which counts 1.146.616 sold units of its Sunday edition in the third quarter of 2023 and about 500 million page visits each month in 2023 on its online news portal *Bild.de* (source: Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e. V. (IVW) (2024)). It is generally known for its polarizing style of reporting characterized by exaggerations, neologisms and an oversimplifying language. In contrast, the two most popular nationwide daily newspapers that can be considered as serious news sources are *Die Süddeutsche Zeitung* (SZ) and the *Frankfurter Allgemeine Zeitung* (FAZ). Each of them counts 454.319 and 498.730 sold units respectively in the third quarter of 2023 and around 55 and 65 million monthly page visits on their news portals *Süddeutsche.de* and *FAZ.net* in 2023 (source: IVW). I use articles from both sources, SZ and FAZ, to increase the sample of "serious" news articles and to cover a larger group of readers in the German population.

I collect SZ and FAZ articles from their respective news archives using a search query that filters articles linked to the topic of inflation in Germany. The queries employ the

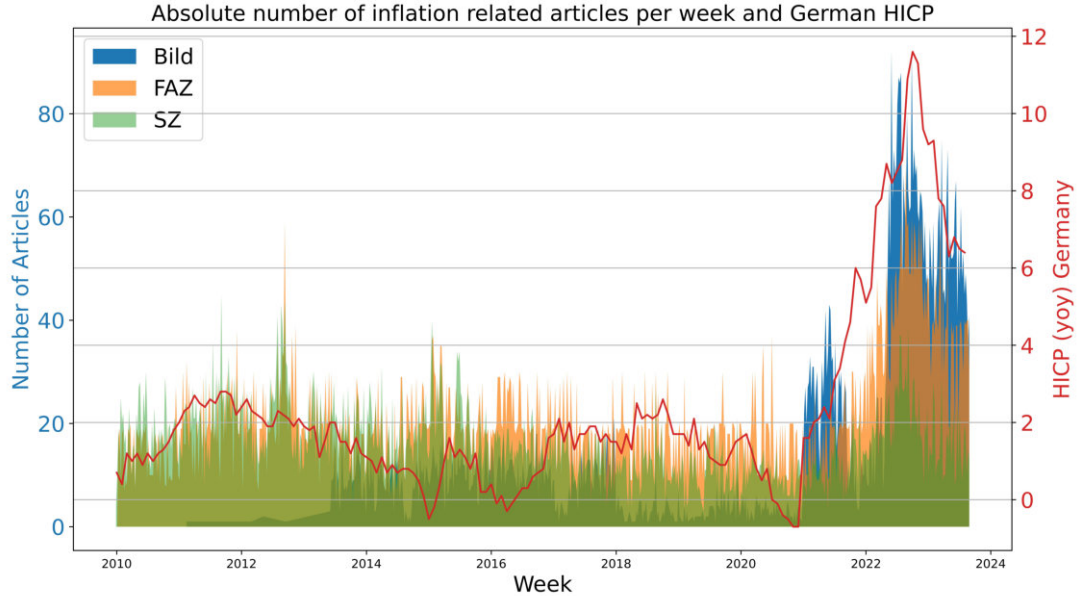


Figure 2: Left axis: weekly absolute numbers of SZ (green), FAZ (orange), and Bild (blue) newspaper articles dealing with inflation in Germany (selection based on classification by ChatGPT). Right axis: monthly year-on-year growth rates of German HICP (red line). Sample span: January 2010 to August 2023.

German versions of words such as *central bank*, *high prices*, *deflation*, *inflation*, *price index*, *wage-price-spiral* or *monetary policy* (the detailed search queries for all newspapers can be found in Appendix A). Both archives comprise print and online articles. Bild newspaper articles are obtained from the online database *LexisNexis*. Since Bild articles are only available until 2017 there, I additionally use the open source platform *mediacloud.org* (mediacloud henceforth) to collect URLs directing to inflation related Bild online articles. I scrape all the collected web pages to obtain the article data. This procedure finally leads to a total data set of 13.365 SZ, 16.157 FAZ, and 16.002 Bild articles published from January 2010 to August 2023. Despite the use of specific search queries, not all articles collected are actually concerned with inflation in Germany. This holds especially true for Bild articles. Therefore, Figure 2 depicts the distribution of articles over time that are actually related to the topic of inflation in Germany against the year-on-year growth rates of the German harmonized index of consumer prices (HICP). At the beginning of the sample, there are only very few Bild articles on inflation which is due to a lack of data availability in the respective online data bases. The number of articles increases a lot starting from 2021. In contrast, for the SZ and FAZ, a fair amount of inflation articles is available throughout the whole sample which starts to pick up in 2021 as well, simultaneously with year-on-year growth rates of German HICP.

2.2 Survey data

To assess the informativeness of the newspaper indicators regarding inflation expectations, I compare them to two separate household expectation surveys, the EC’s *Business and Consumer Survey* (BCS) and the *Consumer Expectations Survey* (CES) conducted by the European Central Bank (ECB). The Former is conducted on a monthly basis in all countries of the European Union. For Germany, the sample dates back to 1985 and includes around 2.000 representative households per survey wave. Participants are surveyed in the first two to three weeks of each month and provide answers to the question “By comparison with the past 12 months, how do you expect that consumer prices will develop in the next 12 months? They will...”, and allows the answers “increase more rapidly” (*PP*), “increase at the same rate” (*P*), “increase at a slower rate”, “stay about the same” (*M*), “fall” (*MM*), or “don’t know.” These qualitative answers get aggregated to an index according to the following formula:

$$\mathbb{E}_t^{\pi_{t+12}} = PP + \frac{P}{2} - \frac{M}{2} - MM \quad (1)$$

The EC provides seasonally adjusted data, dis-aggregated at the following subcategories: age (four bins), education (primary, secondary, and further), gender, income quartiles and occupation status.

As the EC survey data only provides information on the tendency whether households expect inflation to go up or down, I additionally use the ECB’s consumer survey, which runs since April 2020. On a monthly basis, the CES asks between 2.000 and 3.000 participants for their quantitative inflation expectations: “How much higher (lower) do you think prices in general will be 12 months from now in the country you currently live in? Please give your best guess of the change in percentage terms. You can provide a number up to one decimal place.” The survey micro data is available online and provides background information on participant’s age, gender, and education and income level. To correct for outliers I follow Bańkowska et al. (2021) and winsorize all quantitative observations at their 2nd and 98th percentile of their weighted distributions. In a second step, I compute weighted averages for each gender, age, education and income category.

3 Newspaper-based inflation expectation indices

3.1 Applying ChatGPT for sentiment analysis

I extract inflation sentiment signals from the newspaper texts with OpenAI’s ChatGPT-3.5 model (precisely, *gpt-3.5-turbo-1106*), a large language model trained on data up to September 2021. Specifically, I feed each article to the prompt depicted in Figure 3 using OpenAI’s Python API. ChatGPT is asked to provide a sentiment score for each article to capture if the article creates the impression that inflation is likely to in- or decrease. The prompt is designed in a way such that the generated answers are either integers

on a scale from 1 (inflation decreases) to 10 (inflation increases) or the word 'no', if the respective article is actually not about inflation in Germany. I set the model's temperature parameter to 0, which makes it generate a deterministic output when providing identical inputs (Wang et al., 2023). Specifying a system role additionally helps the model to fulfill this task correctly.

```
System: You are an AI that rates inflation sentiment from 1 to 10. You can only respond with a number between 1 and 10.
```

```
User: Provide a sentiment score for inflation in this article on a scale of 1 to 10, where 1 means inflation is likely to decrease, and 10 means inflation is likely to increase. If the article does not address inflation in Germany or the EU, respond with the word 'no': "%title" + "%article"
```

Figure 3: Prompt for generating inflation expectation sentiments with ChatGPT-3.5. "%title" and "%article" are place holders for the single articles and their headlines respectively.

In comparison to simpler methods for text-based sentiment measurement, like the widely-used dictionary approach (see for example Shapiro et al. (2022) or Angelico et al. (2022)), an advantage of ChatGPT is its advanced ability to understand human language, capturing also nuances in the tone of the text. This is especially important when analyzing yellow press articles like those of the Bild newspaper, as they are characterized by a polemic language and constant creation of new expressions. Also, the highly adaptive nature of the Bild language makes creating a Bild suited dictionary impractical as it would require to expand constantly. Table 1 presents examples of Bild expressions and tentative English translations thereof. As shown in Appendix B, ChatGPT is indeed capable of understanding the expressions highlighted in Table 1 and captures the headline's intended tone.

Besides ChatGPT's extended ability to understand human language, the model additionally proves to perform well at analyzing the sentiments of texts. Belal et al. (2023) find that ChatGPT-3.5 significantly outperform two lexicon-based algorithms (VADER and TextBlob) in analyzing the sentiment of amazon product reviews and a soccer tweets data set. Comparing ChatGPT-3.5 to a fine-tuned BERT¹ model, Wang et al. (2023) find that ChatGPT scores equally well on zero-shot sentiment classification tasks and that it even beats the BERT model in recognizing sentiment polarity shifts.

3.2 Index construction

The raw data generated by ChatGPT consists of article-level sentiment scores from 1 to 10, or 'no' if the article does not cover German inflation. Since the frequency of articles is daily, it is feasible to aggregate the data into indices at any higher frequencies. In this analysis, I concentrate on monthly aggregates to align with the monthly frequency of the

¹The BERT (Bidirectional Encoder Representations from Transformers) model is another large language model originally developed by Google (Devlin et al., 2018). Its set of parameters can additionally be fine-tuned on a specific task or dataset.

German Headline	Translation	Date
Kommentar; Schluss mit der Inflations-Heuchelei	Commentary; End to the inflation-hypocrisy .	10/12/2021
Ausgerechnet in der Inflations-Krise!; Deutschlands oberster Geld-Hüter gibt auf	Right in the inflation crisis!; Germany's top money guardian gives up	10/21/2021
Madame Inflation; Luxus-Lagarde macht Sparer und Rentner arm	Madame Inflation; Luxury Lagarde impoverishes savers and retirees	10/30/2021
Nach dem Teuer-Schock ; So einfach schützen Sie Ihr Geld vor Madame Inflation!	After the expensiveness-shock : How to protect your money from Madame inflation!	11/18/2021
Zins-Hammer! EZB erhöht auf 1,25 Prozent	Interest rate hammer! ECB increases to 1.25 percent	09/08/2022

Table 1: Examples of Bild typical language and their use of neologisms in the context of inflation in Germany and the euro zone.

household expectation surveys. In line with Angelico et al. (2022), I use multiple methods to aggregate the single scores to sentiment indices for each newspaper separately and only include those articles, which ChatGPT deems to be related to inflation in Germany:

1. **Index #1:** I take the mean over all articles published per month to obtain a monthly sentiment index.
2. **Index #2:** First, I obtain mean scores per day to then compute moving averages (MA) over a window of 10, 30 or 60 days (MA_10, MA_30 & MA_60). I finally aggregate them to monthly frequency by taking the average for each month.
3. **Index #3:** Analogously to Angelico et al. (2022), I compute exponentially weighted averages over the daily mean scores using three alternative smoothing parameters: 0.1, 0.3 and the one that minimizes the sum-of-squared in-sample forecast errors, which is close to 0.1 (indices are denoted by EWM_01, EWM_03 & EWM_opt).
4. **Index #4:** I reproduce all indices #1 to #3 including only articles that score 5 or above on ChatGPT's inflation sentiment scale.

I choose to construct the indices also on a sub-sample based on articles with higher sentiment scores as it might be the case that newspapers rather report on increasing inflation than on stable, low price levels. The increase in the number of inflation related articles shown in Figure 2 as well as the worsening of the average sentiment depicted in Figure 4 is supporting evidence for this hypothesis. Therefore, ChatGPT might also be more reliable when assigning high sentiment scores. Similarly, Weber et al. (2023) find that households also pay more attention to inflation, when inflation is increasing, therefore an index based on higher scores might be more successful in capturing their inflation expectations.

Figure 4 exemplary shows one index from each of the categories #1 to #3 as well as their versions based on articles with sentiment scores greater than 4 for each of the newspapers separately and the FAZ and SZ articles data combined. Overall, the EWM_01 and MA_60 indices seem to be less noisy than the mean indices. While the indices based on FAZ and SZ data are very well aligned with each other, the Bild inflation indices significantly differ. In all depicted cases, the Bild inflation index reaches higher sentiment scores than the other two newspaper, which is especially visible in the years after 2019. Due to the scarcity of Bild articles at the beginning of the sample, the Bild sentiment index is often inconclusive for the period before 2014. Therefore, I will truncate the sample at 2014 for the Bild newspaper in subsequent analyses. Comparing the newspaper indices to actual year-on-year growth of German HICP reveals some correlation over time, with the newspaper indicators beginning to align with inflation trends at the end of 2021. When considering only articles with a sentiment score above 4, it becomes evident that the disparity between Bild and the FAZ and SZ newspapers is driven by Bild’s abundance of articles suggesting high or increasing inflation.

4 Informativeness of the newspaper indices

4.1 Heterogeneous correlations

In a first step, I compare the sample correlations between the different newspaper inflation sentiment indices and the different subgroups of the EC and ECB household survey expectations. This reveals first heterogeneity in how strongly which newspaper index co-moves with which subgroup of the households.

In general one can say that in the case of EC survey data, when considering the full sample, the Bild achieves the highest sample correlations (ranging between 0.53 and 0.82) among the three newspapers and that indices based on scores higher than 4 produce a lower correlation than those based on all ChatGPT sentiment scores. While the Bild indices being only weakly correlated with German HICP, FAZ and SZ indices attain a higher correlation with HICP than the survey expectations themselves. The first two of these results are mainly driven by the sample period from 2014 to 2019—when Germany experienced rather low and stable levels of inflation. Starting from 2020, the newspaper indices suddenly have a stronger correlation with HICP than the survey expectations. In this period, there is also a lower dispersion in the correlations between the different sentiment score aggregation methods what might stem from the fact, that there is a higher share of articles with high sentiment scores in these times.

Considering the sample period from 2014 to August 2023, there is in fact some heterogeneity between the correlation of the newspaper indices and the different household groups of the EC survey. The correlation heat maps in Figure 5 show that for the Bild, across education levels, the highest level is the least correlated. Meanwhile, the differ-

ences in correlation within the other sub-categories are less strong. In contrast, the FAZ shows the largest differences in correlations within sub-categories for the age criterion, being most correlated with younger cohorts expectations. This changes when cutting the sample at March 2020. Then, the FAZ and SZ indices achieve the highest correlation with the expectations of survey participants with the highest education level. For the Bild indices, the correlation pattern is rather consistent over time.

When comparing the newspaper indices to the ECB survey expectations, overall, the correlation with the Bild indices is lowest among the three newspapers with a generally low heterogeneity across household sub-categories. However, as distinguishable in Figure 6, the correlation with the FAZ indices increases with the level of education and income. The same holds true for the SZ indices, even though they generally reach lower correlations.

4.2 Heterogeneous forecasting powers

In a second step, I run an out-of-sample forecasting exercise and compare the ability of the different newspaper indices to predict the survey expectations of the various groups of households. As benchmark, I estimate an autoregressive model, $AR(p)$, for each of the sub-aggregated survey expectation measures and each survey (ECB and EC) separately, determining the optimal lag order p for each model and forecast iteration with the BIC criterion. For the EC data, the in-sample ends in 2016M12, translating into a training sample length of 84 months for SZ and FAZ newspapers and 36 months for the Bild, due to the shorter data availability of the latter. When predicting ECB survey data, I use an in-sample of 12 months so that the out-of-sample starts in 2021M04. Then, I augment each benchmark model with one of the 14 different inflation expectation indices for each newspaper separately and also a combined version of FAZ and SZ data. To be conservative as for both surveys the participants are surveyed in the first weeks of a month, all newspaper indices enter with one lag. The forecast horizon for EC data is 6 months and for the ECB survey 3, due to its short sample length. All horizons are predicted recursively.

I compare the forecasting powers by calculating the root mean squared error (RMSE) for each model. I divide the augmented model's RMSE by the benchmark's RMSE such that a value smaller than one indicates an improvement in forecast accuracy when including the newspaper indices. In general, the RMSE ratios are very similar across the 14 aggregation methods for each newspaper with point wise differences. Therefore, and to make the results more tractable, I only report their average values here².

Table 2 presents the relative RMSEs based on EC survey data. One can see that in most cases, including the newspaper expectation indices leads to an improvement against the benchmark. There is also sizable heterogeneity between the single newspapers. While

²The single RMSE ratios for each aggregation method of each newspaper and survey can be found in the online appendix.

the Bild indices perform better in predicting expectations of lower than higher income quartiles, FAZ and SZ also carry informativeness for the third income quartile. Contrasting the newspaper’s performance in predicting expectations across different education levels, the Bild performs best in predicting the lowest, while FAZ and SZ perform best in predicting the highest education level. For longer horizons, FAZ and SZ fail to improve predictions of gender-specific expectations and expectations of age class 3 and 4 (age 50 to 65+). Here, at least beginning from horizon 2, the Bild indices can lead to improvements. Nonetheless, FAZ and SZ carry a greater informativeness advantage when predicting inflation in Germany, compared to the Bild.

Table 3 shows the results of the forecasting exercise based on ECB survey data. Almost all relative RMSEs are smaller than one which means that there are information gains by including the newspaper indices. For the Bild, there is not much heterogeneity across the household sub-categories. This is different especially to the FAZ and to some extent the SZ newspaper. Both perform best in predicting the expectations of the highest income and education level and the age 35 to 49 cohort. All three newspapers are more informative in predicting male versus female inflation expectations.

These results already suggest, that there is a pattern such that the Bild inflation expectation indicators are better predictors of lower income and lower education level household’s expectations while the sentiment conveyed in newspapers such as the FAZ and SZ rather resonate with higher income and more educated households as well as actual inflation figures. To shed some more light, I also compare the RMSEs of the two types of newspaper against each other. For this purpose, I divide the FAZ and SZ RMSEs by the Bild RMSEs, such that values below 1 identify an information advantage of the FAZ and SZ indices. Table 4 and 5 report that in almost all cases, the FAZ and SZ indices outperform the Bild indices in predicting survey expectations and inflation in Germany. One reason for this observation might be the different training sample lengths of Bild and FAZ and SZ data. Nonetheless, one can detect some tendencies: for EC data (Table 4), the Bild is relatively better in predicting lower income quartile and lower education level, as well as middle aged households. The same pattern emerges for the ECB survey based results reported in Table 5.

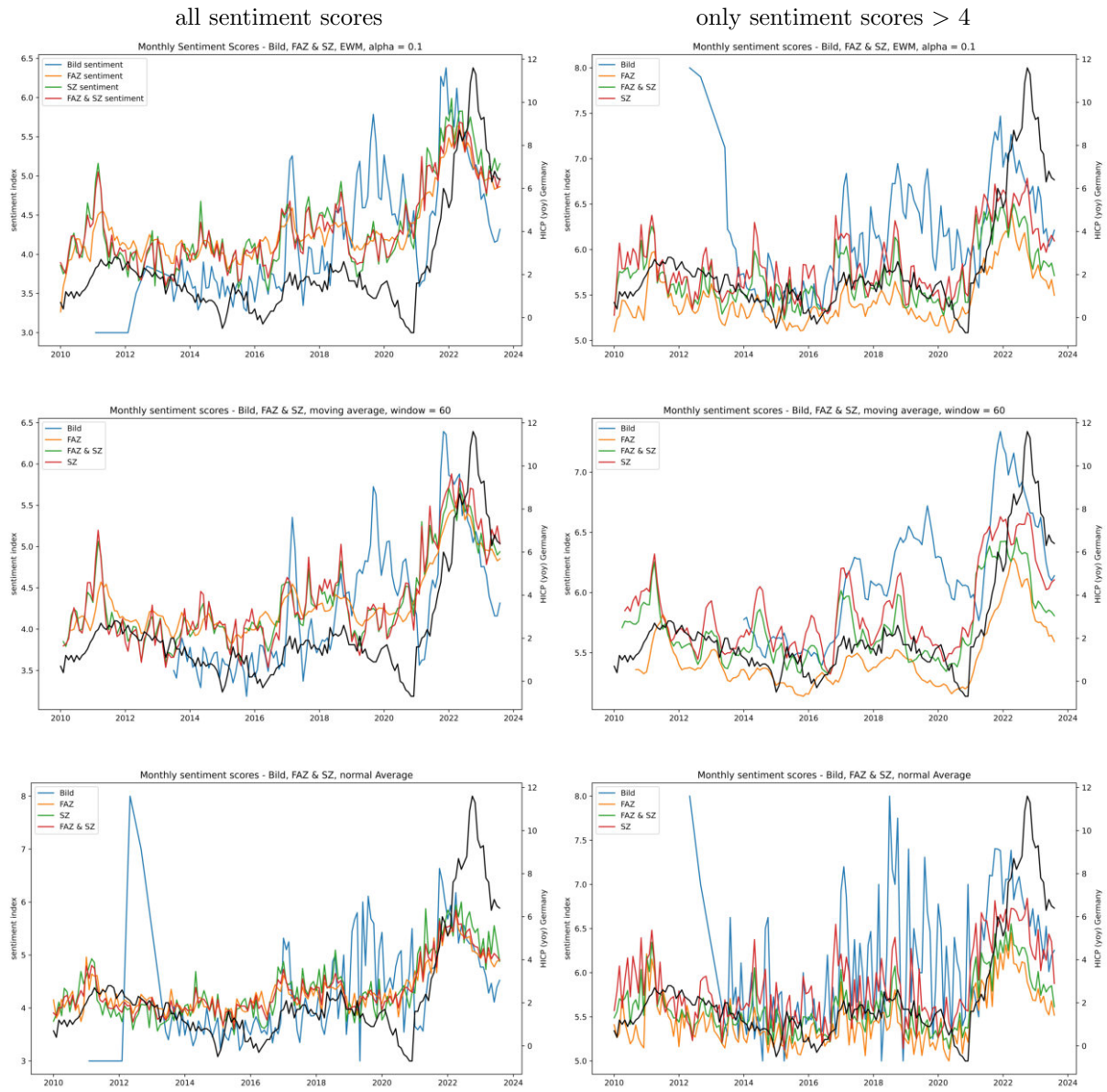


Figure 4: Left column: shows the EWM.01, MA.60 and mean (top to bottom) sentiment indices (left axis) for each newspaper on a monthly frequency against German year-on-year HICP growth rates (black line, right axis). Right column: corresponding versions of the indices using only articles with sentiment scores above 4.

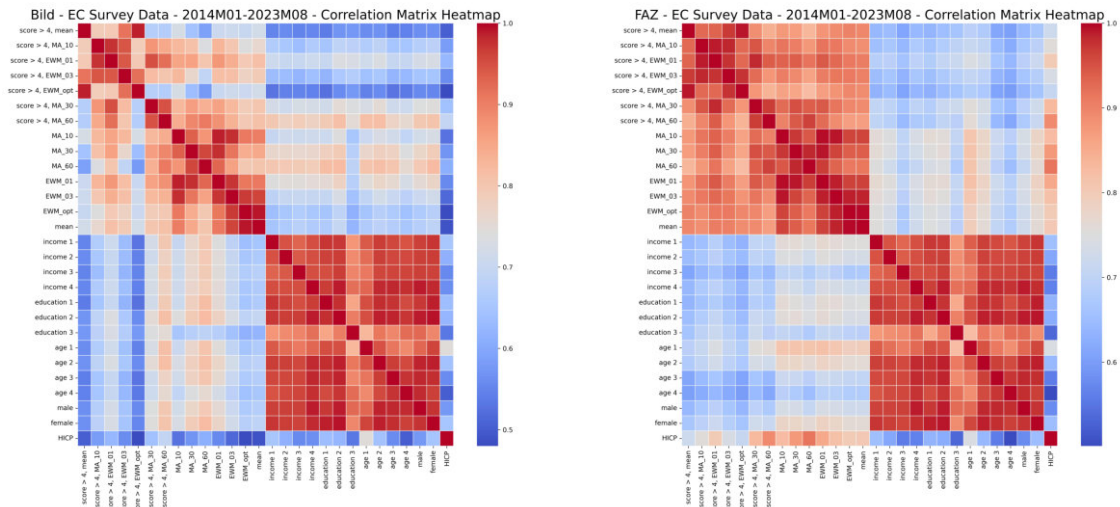


Figure 5: Sample correlations between EC data and Bild (left) and FAZ (right) inflation indices based on the whole sample period (2014M01-2023M08). Red indicates high and blue low correlations.

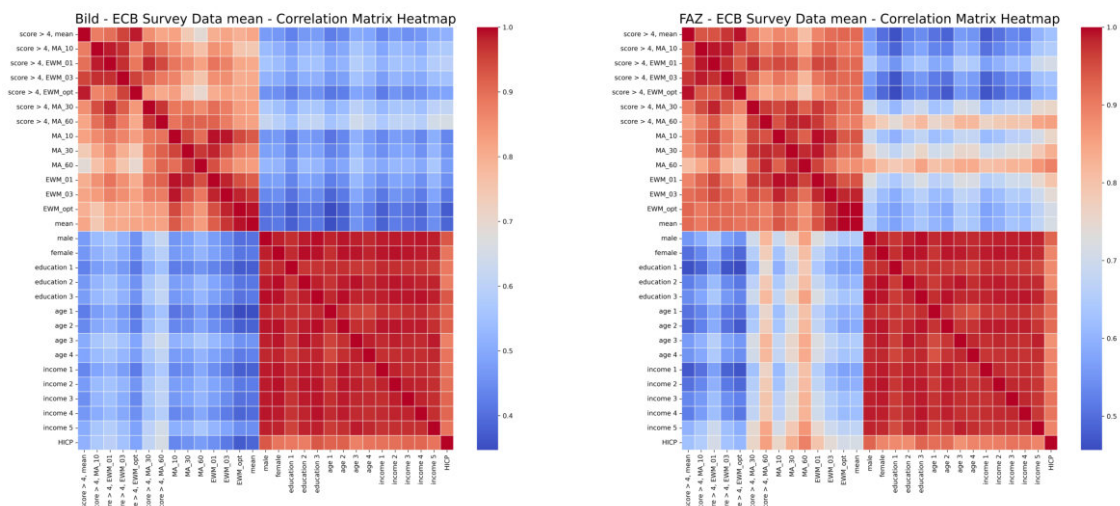


Figure 6: Sample correlations between ECB survey data and Bild (left) and FAZ (right) inflation indices based on the whole sample period (2020M04-2023M08). Red indicates high and blue low correlations.

News	income 1	income 2	income 3	income 4	edu 1	edu 2	edu 3	age 1	age 2	age 3	age 4	male	female	HICP
Bild	0.99	0.99	1.01	1.03	1.00	1.00	1.02	1.00	1.02	1.00	1.02	1.00	1.01	0.99
	0.96	0.97	1.00	1.01	0.98	0.98	0.98	0.94	0.99	0.99	1.00	0.98	0.99	0.98
	0.94	0.95	0.98	0.99	0.96	0.97	0.96	0.91	0.98	0.97	0.99	0.97	0.97	0.97
	0.91	0.92	0.96	0.97	0.94	0.95	0.93	0.88	0.96	0.96	0.99	0.97	0.95	0.95
	0.89	0.89	0.94	0.95	0.91	0.93	0.90	0.85	0.94	0.94	0.97	0.95	0.93	0.94
	0.86	0.86	0.91	0.92	0.89	0.91	0.87	0.83	0.91	0.92	0.96	0.93	0.91	0.93
FAZ	0.94	0.92	0.98	0.99	0.98	1.00	0.95	0.90	0.99	1.00	1.00	0.99	1.00	0.88
	0.95	0.92	0.96	1.00	0.99	1.01	0.95	0.86	0.97	1.01	1.02	1.01	1.01	0.81
	0.94	0.92	0.96	1.00	0.99	1.02	0.94	0.84	0.97	1.02	1.03	1.01	1.01	0.77
	0.91	0.92	0.94	1.01	0.99	1.02	0.94	0.82	0.96	1.03	1.03	1.01	1.01	0.71
	0.90	0.91	0.94	1.01	0.98	1.01	0.93	0.80	0.96	1.03	1.03	1.01	1.00	0.66
	0.90	0.92	0.94	1.02	0.99	1.01	0.93	0.80	0.96	1.03	1.03	1.02	1.00	0.63
SZ	0.95	0.93	0.99	0.99	0.98	1.01	0.96	0.91	0.99	1.00	0.99	0.99	1.00	0.90
	0.96	0.92	0.96	0.99	0.98	1.01	0.95	0.86	0.97	1.00	1.01	1.00	1.00	0.86
	0.95	0.92	0.96	1.00	0.98	1.01	0.95	0.84	0.97	1.01	1.02	1.00	1.01	0.83
	0.93	0.92	0.95	1.01	0.98	1.02	0.95	0.82	0.97	1.02	1.03	1.01	1.01	0.80
	0.93	0.91	0.95	1.02	0.98	1.02	0.94	0.81	0.97	1.03	1.03	1.01	1.01	0.77
	0.94	0.92	0.94	1.02	0.99	1.01	0.95	0.80	0.98	1.03	1.03	1.02	1.01	0.75
FAZ & SZ	0.94	0.92	0.98	0.99	0.98	1.00	0.96	0.90	0.98	0.99	0.99	0.99	1.00	0.89
	0.95	0.91	0.95	0.98	0.98	0.99	0.93	0.85	0.96	1.00	1.00	0.99	0.99	0.84
	0.93	0.91	0.94	0.99	0.98	0.99	0.92	0.83	0.95	1.00	1.01	0.99	1.00	0.80
	0.91	0.91	0.92	1.00	0.98	0.99	0.91	0.81	0.95	1.01	1.01	0.99	0.99	0.75
	0.90	0.90	0.92	1.00	0.97	0.99	0.90	0.79	0.94	1.01	1.01	0.99	0.99	0.72
	0.90	0.91	0.91	1.00	0.98	0.98	0.90	0.78	0.94	1.01	1.01	0.99	0.98	0.70

Table 2: Relative root mean squared errors of the averaged newspaper indicators of the FAZ, SZ, Bild and FAZ & SZ combined compared to EC survey data. Each row per newspaper stands for a different forecast horizon, starting with $h = 1$ in the first row.

News	age 1	age 2	age 3	age 4	income 1	income 2	income 3	income 4	income 5	edu 1	edu 2	edu 3	male	female
Bild	0.89	0.92	0.96	0.91	1.12	0.97	0.97	0.91	0.93	0.96	0.93	0.90	0.86	0.99
	0.85	0.92	0.91	0.90	1.13	0.93	0.93	0.89	0.92	0.95	0.88	0.89	0.83	0.94
	0.82	0.88	0.88	0.90	1.18	0.90	0.92	0.86	0.89	0.94	0.85	0.85	0.79	0.91
FAZ	0.84	0.77	0.95	0.78	0.86	0.85	0.90	0.83	0.72	0.86	0.83	0.73	0.76	0.86
	0.78	0.73	0.82	0.75	0.80	0.80	0.83	0.76	0.65	0.84	0.76	0.69	0.70	0.80
	0.73	0.67	0.76	0.73	0.75	0.75	0.79	0.70	0.56	0.82	0.71	0.61	0.63	0.74
SZ	0.88	0.85	0.97	0.88	0.96	0.90	0.91	0.85	0.81	0.94	0.90	0.80	0.82	0.93
	0.84	0.84	0.88	0.86	0.93	0.87	0.86	0.83	0.77	0.89	0.86	0.78	0.79	0.88
	0.81	0.79	0.82	0.83	0.91	0.84	0.84	0.80	0.70	0.88	0.83	0.73	0.74	0.85
FAZ & SZ	0.88	0.81	0.92	0.86	0.96	0.87	0.89	0.82	0.80	0.91	0.86	0.77	0.79	0.88
	0.84	0.79	0.84	0.83	0.93	0.84	0.84	0.79	0.76	0.88	0.82	0.76	0.76	0.84
	0.81	0.75	0.78	0.81	0.91	0.80	0.82	0.75	0.70	0.86	0.79	0.71	0.71	0.81

Table 3: Relative root mean squared errors of the averaged newspaper indicators of the FAZ, SZ, Bild and FAZ & SZ combined compared to ECB survey data. Each row per newspaper stands for a different forecast horizon, starting with $h = 1$ in the first row.

News	income 1	income 2	income 3	income 4	edu 1	edu 2	edu 3	age 1	age 2	age 3	age 4	male	female	HICP
FAZ vs. Bild	0.90	0.88	0.92	0.90	0.94	0.93	0.85	0.84	0.95	0.93	0.93	0.93	0.92	0.85
	0.93	0.89	0.90	0.92	0.96	0.95	0.86	0.84	0.92	0.95	0.95	0.95	0.94	0.81
	0.93	0.90	0.92	0.94	0.97	0.96	0.87	0.85	0.92	0.96	0.96	0.95	0.96	0.77
	0.92	0.92	0.92	0.94	0.98	0.96	0.87	0.87	0.92	0.98	0.96	0.95	0.97	0.72
	0.93	0.94	0.92	0.95	0.99	0.96	0.88	0.88	0.92	0.99	0.97	0.95	0.98	0.68
	0.95	0.97	0.93	0.97	1.02	0.97	0.90	0.89	0.94	1.00	0.98	0.97	0.99	0.65
SZ vs. Bild	0.91	0.88	0.92	0.90	0.94	0.94	0.86	0.85	0.95	0.93	0.93	0.93	0.93	0.87
	0.94	0.88	0.90	0.91	0.95	0.95	0.86	0.84	0.92	0.94	0.94	0.94	0.94	0.86
	0.94	0.90	0.92	0.93	0.96	0.96	0.88	0.85	0.92	0.96	0.95	0.94	0.96	0.84
	0.94	0.92	0.92	0.94	0.98	0.96	0.88	0.88	0.93	0.98	0.96	0.94	0.97	0.82
	0.96	0.94	0.93	0.95	0.99	0.97	0.89	0.88	0.94	0.99	0.96	0.95	0.98	0.80
	0.99	0.97	0.93	0.97	1.02	0.97	0.92	0.89	0.96	1.00	0.97	0.97	1.00	0.78

Table 4: Relative root mean squared errors of the averaged newspaper indicators of the FAZ and SZ compared to Bild RMSEs, based on EC survey data. Each row per newspaper stands for a different forecast horizon, starting with $h = 1$ in the first row.

News	age 1	age 2	age 3	age 4	income 1	income 2	income 3	income 4	income 5	edu 1	edu 2	edu 3	male	female
FAZ vs. Bild	0.94	0.84	0.99	0.86	0.77	0.88	0.93	0.91	0.77	0.90	0.89	0.81	0.89	0.87
	0.91	0.80	0.91	0.83	0.72	0.86	0.89	0.85	0.71	0.88	0.86	0.77	0.84	0.85
	0.89	0.76	0.86	0.81	0.66	0.83	0.86	0.81	0.63	0.87	0.84	0.72	0.80	0.81
SZ vs. Bild	0.98	0.93	1.01	0.97	0.86	0.93	0.93	0.94	0.87	0.98	0.96	0.88	0.95	0.95
	0.98	0.92	0.97	0.95	0.84	0.94	0.92	0.93	0.84	0.94	0.98	0.88	0.95	0.94
	0.99	0.90	0.93	0.92	0.79	0.93	0.91	0.92	0.80	0.94	0.98	0.86	0.95	0.93

Table 5: Relative root mean squared errors of the averaged newspaper indicators of the FAZ and SZ compared to Bild RMSEs, based on ECB survey data. Each row per newspaper stands for a different forecast horizon, starting with $h = 1$ in the first row.

Another way to evaluate the forecast performance of competing models against each other and further identify heterogeneity in the predictability of household expectations by the different newspaper indices is by computing the cumulative sum of squared forecasting errors differences (CSSED), defined as:

$$CSSED_{m,t} = \sum_{t=1}^T (\hat{e}_{bm,t}^2 - \hat{e}_{m,t}^2), \quad (2)$$

where \hat{e}_{bm} and \hat{e}_m correspondingly denote the forecasting errors of the benchmark and newspaper indices augmented models in the out-of-sample. Similar to the previous analysis, I average the forecast errors across the 14 aggregation methods for the results I report here and present all single index results in the online appendix. If the CSSED falls below 0 at a particular juncture during out-of-sample prediction, it indicates that when we divide the in-sample and evaluation samples at that juncture, the benchmark model performs better, displaying a lower RMSE compared to the competing model. Conversely, if the CSSED exceeds 0, then the competing model surpasses the benchmark.

Figure 7 shows the one month ahead CSSEDs of the averaged newspaper indices targeting EC survey data. Again, one can see heterogeneity in which newspaper index is more useful in predicting the inflation expectations of different household groups. In general, the Bild index appears to be more informative for most groups of households than the other two newspapers as the CSSED exceeds zero already early on in the out-of-sample, while for FAZ and SZ this is mainly the case starting in 2021. The Bild index generates the largest CSSED for male, lower income and education, as well as middle aged respondent's expectations and is at par with the benchmark when predicting HICP. In contrast, for the FAZ and SZ indices the CSSEDs are especially large when predicting inflation expectations of higher educated participants, or those belonging to the second or third income quartile.

Figure 8 presents the results of this exercise when predicting ECB survey expectations. Besides a few outlier (for example, Bild's poor performance when forecasting expectations of the first income quartile), there seems to be less heterogeneity across the three newspaper indices. They all start to exceed the zero in 2022. This might be due to the short out-of-sample length. For the Bild, the CSSEDs of the different sub-categories are very close to each other, but the largest values are reached when targeting younger and lower educated respondents. The FAZ and SZ CSSEDs are also quite similar, but the indices perform best when targeting the highest income quartile and worst when predicting female and age 30-49 participant's expectations.

5 Monetary policy shocks and news sentiments

To explore a possible source of the heterogeneity between the newspaper indices, I estimate the newspaper's sentiment reaction after a monetary policy shock. Therefore, I use a local

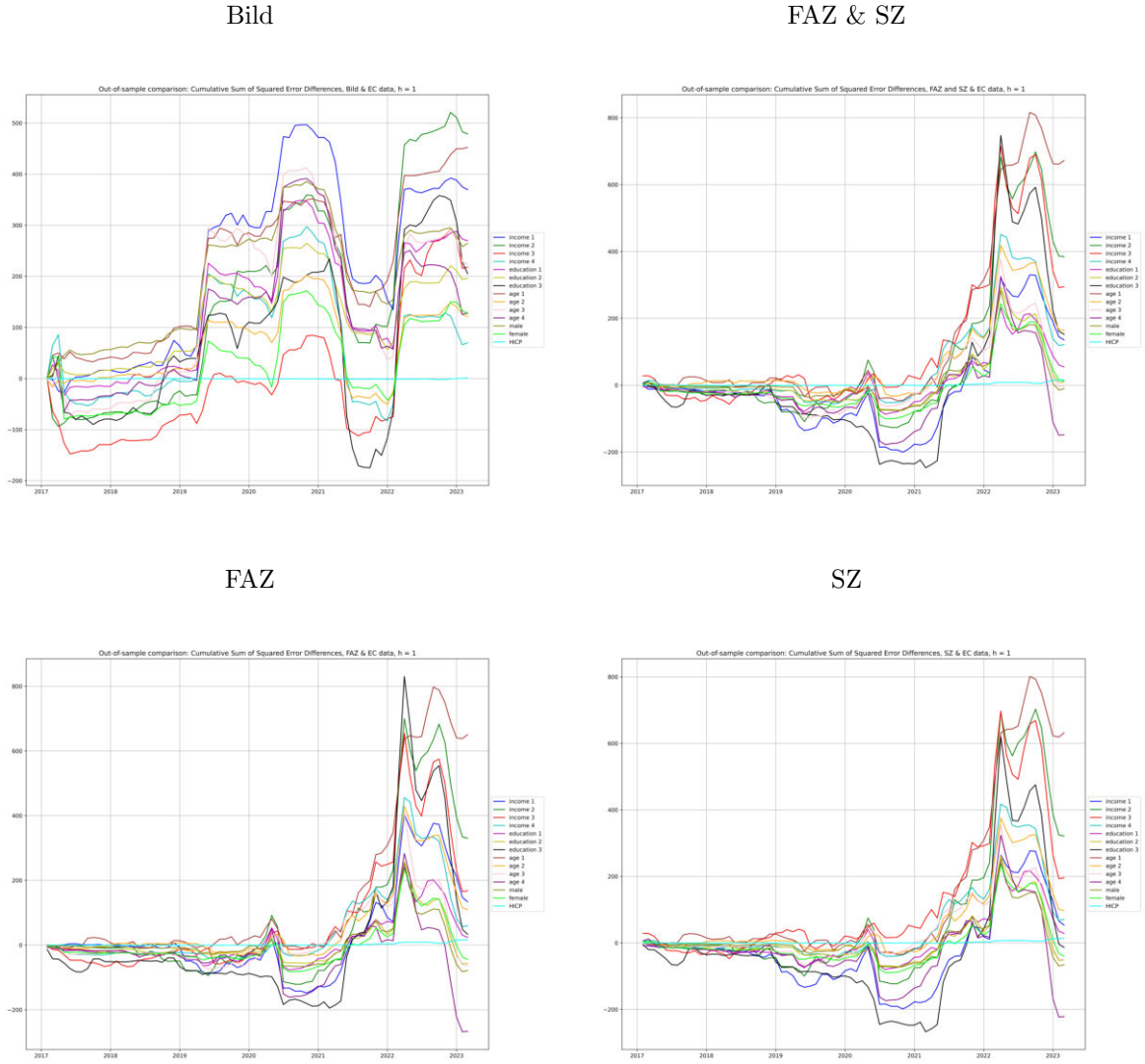


Figure 7: Cumulative sum of squared forecasting errors differences for the one month ahead out-of-sample forecasts of the averaged Bild, FAZ, SZ, and FAZ and SZ data combined newspaper expectation index. If the CSSED is above 0, the competing model fares better than the benchmark $AR(p)$. The target variables are the EC survey’s sub-categories.

projections model as introduced by Jordà (2005):

$$y_{t+h} = \mu_h + \beta_h shock_t + \Delta_h \sum_{l=0}^n W_{t-l} + u_{t+h}. \quad (3)$$

I estimate the model at daily frequency for each horizon h separately. The model includes a horizon-specific constant μ_h and the dependent variable y_t is in this case the log differences of one of the three newspaper sentiment indices. Specifically and similar to Born et al. (2023), I use the 7-day moving average of each newspaper’s daily mean sentiment index to smooth out the days without any news reporting. The variable $shock_t$ is a shock series of ECB monetary policy surprises generated as in Jarociński and Karadi (2020) and provided online by Marek Jarociński³. For estimating euro area monetary policy shocks they rely on the *Euro Area Monetary Policy Event-Study Database* devel-

³<https://marekjarocinski.github.io/jkshocks/jkshocks.html>

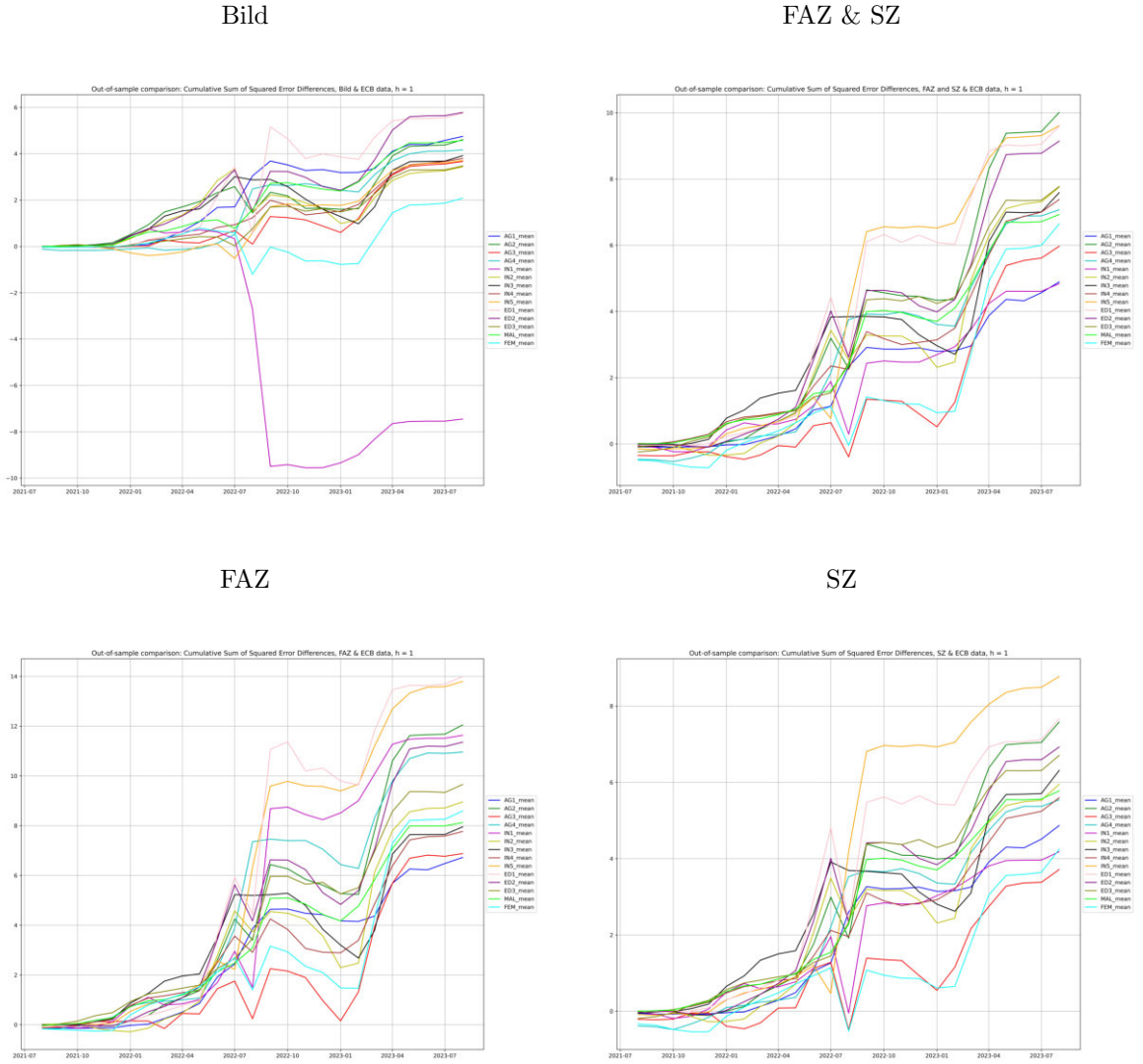


Figure 8: Cumulative sum of squared forecasting errors differences for the one month ahead out-of-sample forecasts of the averaged Bild, FAZ, SZ, and FAZ and SZ data combined newspaper expectation index. If the CSSED is above 0, the competing model fares better than the benchmark $AR(p)$. The target variables are the ECB survey’s sub-categories.

oped by Altavilla et al. (2019) which contains intra-day asset price changes around ECB policy announcements. The shock series is available in daily frequency, even though at maximum one monetary policy decision is taken per month. The controls W_t consist of the contemporaneous value and $n = 14$ lags (i.e., days) of the respective 7-day moving average newspaper sentiment index and German HICP. As inflation is only available at monthly frequency it enters the regression as a series that takes the same inflation value of a respective month for each day of that month. The model is estimated on a sample spanning from 2014M01 to 2023M06, due to the limited data availability of Bild articles at the beginning of the sample and the availability of the shock series. I follow Montiel Olea and Plagborg-Møller (2021) and compute Eicker-Huber-White (EHW) heteroskedasticity-robust standard errors which are sufficient when including lags of the dependent variable

as controls. Finally, the impulse responses are given by the β_h coefficients for each horizon separately.

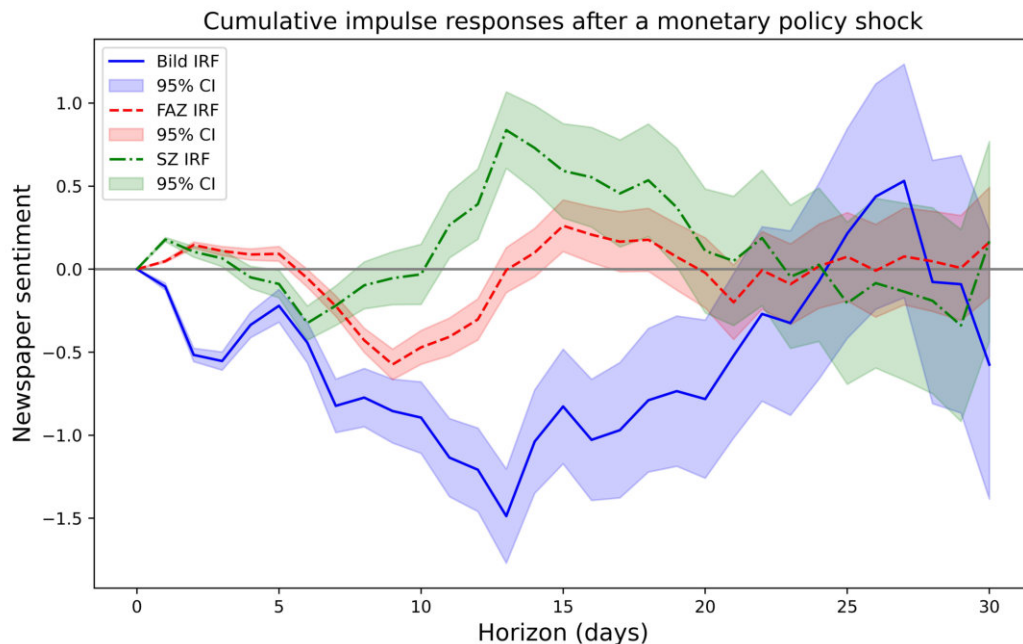


Figure 9: Impulse responses of the three newspaper sentiment indices (Bild in blue, solid line, FAZ in red, dashed line, and SZ in green, dotted-dashed line) to a contractionary monetary policy shock on a 30-days horizon. Shaded areas are 95% confidence intervals computed with EHW standard errors.

Figure 9 shows the resulting cumulative impulse response functions over a horizon of 30 days. One can see that in the first days after a contractionary monetary policy shock, the Bild sentiment directly starts to cool down, while the FAZ and SZ sentiment significantly decreases only after three to five days. In contrast to the response of Bild sentiment, the responses of the SZ and FAZ look quite similar, even though the FAZ sentiment seems to be lagging behind the SZ one for a few days.

All in all, the signals towards inflation in the FAZ and SZ newspaper seem to be less strongly affected by a monetary policy surprise than those of the Bild newspaper. One possible reason for that could be the higher volatility in the Bild sentiment compared to the other two newspapers. Looking at Figure 10, which shows the 7-day moving average newspaper indices against the ECB’s monetary policy meeting dates (vertical lines) during the high inflation period of 2021 until the end of the sample, it seems like the Bild signals increasingly higher inflation right before the ECB meetings and then rapidly reverses its stance. The opposite we can observe for the SZ and FAZ (SZ being more volatile than FAZ), which rarely peak around the monetary policy meetings and rather right between two meetings. On the one hand it can be a good sign, that monetary policy is actually able to cool down yellow press newspaper’s sentiment. On the other side, if this is really driven by a built up in negative sentiment right before the policy meetings, it would be

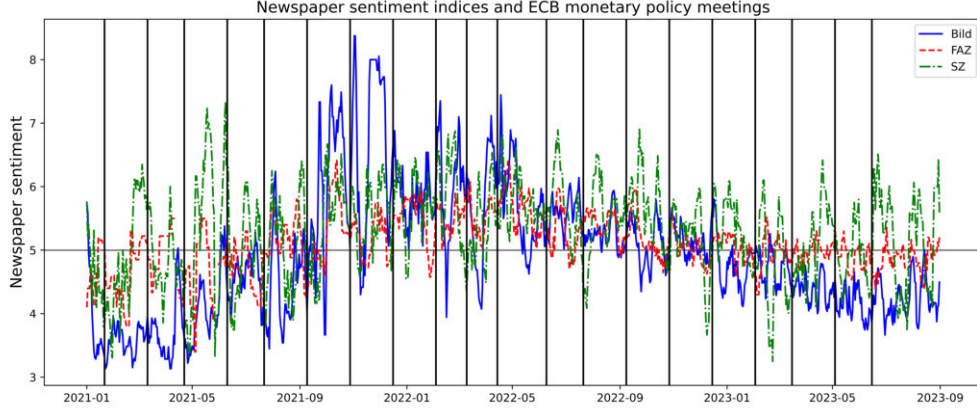


Figure 10: ECB monetary policy meetings (vertical lines) plotted against the three newspaper 7-day moving average sentiment indices (Bild in blue, solid line, FAZ in red, dashed line, and SZ in green, dotted-dashed line), 2021M01-2023M08.

advisable for the ECB to try to moderate the public discourse to prevent exaggerated counterproductive mood-making.

6 Household’s reaction to news sentiment shocks

In a second step, I investigate whether the heterogeneous informativeness of the three newspaper indicators roots in their potential ability to affect household expectations differently. To this end, I closely follow Shapiro et al. (2022)’s modelling strategy using local projections, but make some adaptations to this specific use case. Shapiro et al. (2022) estimate the effect of a news sentiment shock on measures of economic activity as the component of the news sentiment index that is *”orthogonal to current and 4 lags of economic activity as well as 4 lags of itself.”* In contrast, the dependent variable in my model is one of the 13 sub-categories of the EC survey’s household inflation expectations, but I also estimate their impulse responses as their component which is orthogonal to the current value and $n = 6$ lags of the news sentiment index and indicators of economic activity, respectively. I perform this exercise only on EC data, due to its larger sample coverage, compared to the ECB survey. A separate regression is run for each forecast horizon $h = 1, \dots, 6$ months:

$$y_{t+h}^i = \alpha_h^i + \beta_h^i NSI_t + \mathbf{\Gamma}_h^i \sum_{l=0}^n W_{t-l} + \sum_{l=1}^n \delta_{h,l}^i NSI_{t-l} + u_{t+h}^i, \quad (4)$$

where y^i is one of the $i = 1, \dots, 13$ survey measures, α_h^i a horizon specific constant, and NSI denotes one of the four newspaper sentiment indices (Bild, FAZ, SZ and FAZ & SZ combined) as the average across all ways of aggregation excluding those that only take articles with sentiment scores greater than 4 into account. Both, the survey and sentiment measures are expressed in first differences. The set of controls W is also oriented towards Shapiro et al. (2022) and contains the contemporaneous value and 6 lags of the

German industrial production index excluding energy and construction in log-differences (as a monthly proxy for output), year-on-year growth rates of German HICP, and first differences of a measure of consumer confidence as monthly an indicator of consumption. The first two series are obtained from the Federal Statistical Office of Germany while the latter can be retrieved from Eurostat and is compiled by the EC. The impulse responses are then given by estimates of the β_h^i coefficients which I report with 95% confidence bands constructed with Newey-West autocorrelation and heteroskedasticity robust standard errors. I estimate this model on 2014M01 to 2023M08 data with the availability of Bild sentiment scores being the limiting factor.

Figures 11 to 14 show the resulting cumulative impulse response functions to shocks in Bild, FAZ, and SZ (and the latter two combined) inflation sentiments, each figure representing one of the four EC survey sub-categories education, income and age level, and gender. In general, the heterogeneity in the responses is less pronounced for shocks to SZ and Bild sentiment and most significant for the FAZ. Nonetheless, there are some trends distinguishable. Figure 11 shows that FAZ and SZ have the strongest impact on the survey expectations of more educated households (education further than secondary school) while responses of education level 1 and 2 households react less and are mostly insignificantly different from each other. Impulse responses after shock in Bild sentiment are not significantly different from each other and are very close to a zero response.

When dis-aggregating among income quartiles (Figure 12), in the case of the Bild, the second lowest income group is most responsive and significantly different from zero in the first four months after impact, while the response of expectations in the highest income group is even negative most of the time. This might reflect the circumstance that shocks in Bild sentiment do not necessarily reach wealthier households and explain some of the heterogeneity in household expectations. At the same time, shocks to FAZ and SZ sentiment induce the smallest response in the lowest income group (in the case of the SZ it is even mostly insignificant from zero), while the effects are most pronounced in the second and third income quartile.

There is no clear heterogeneity in responses dis-aggregated across age after a shock to SZ and Bild sentiment (see Figure 13), except for the FAZ, where the third age group (age 50 to 64) is the least responsive. Finally, Figure 14 shows that shocks to different newspapers basically do not affect female or male survey participant’s expectations differently. Only the overall effect differs which is largest for the FAZ and smallest or mostly insignificant for the Bild.

7 Conclusion

In this paper, I extract inflation sentiments on the article level from three major daily German newspapers—the Bild, FAZ and SZ, using the LLM ChatGPT-3.5 to construct inflation indices out of them. I then analyze the indices’ informativeness regarding dif-

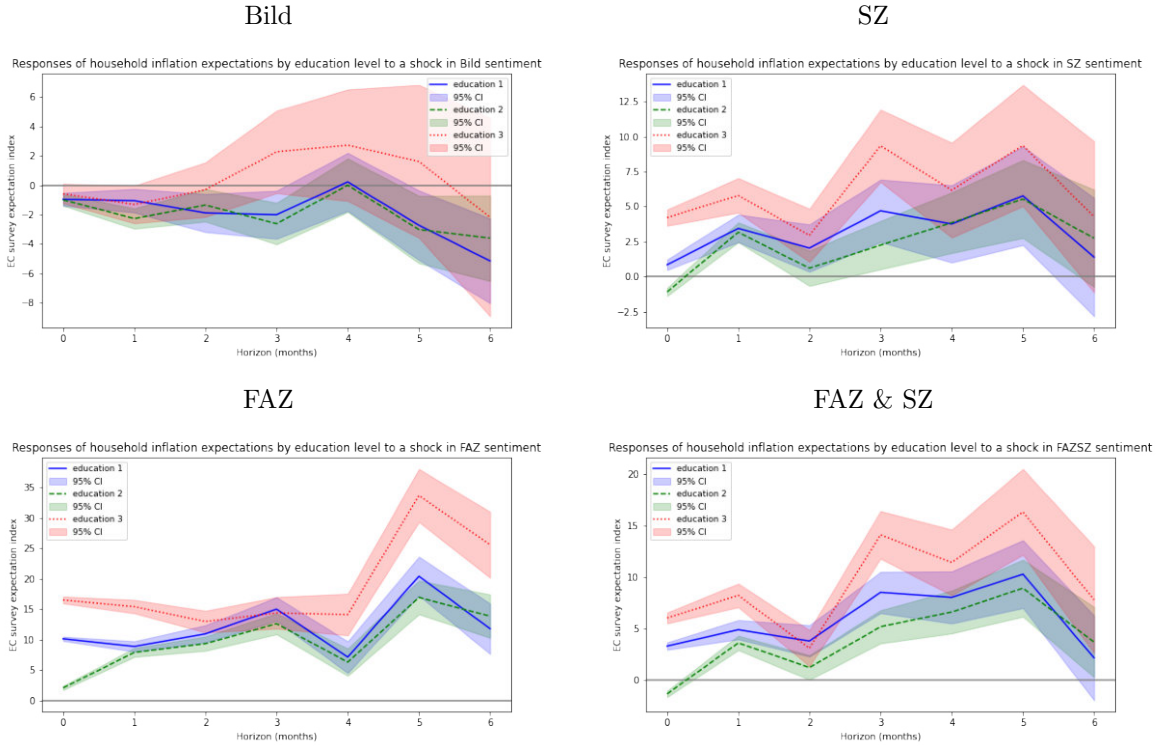


Figure 11: Cumulative impulse responses of EC survey expectations dis-aggregated across education levels to a shock in newspaper sentiment. Shaded areas represent 95% confidence bands constructed with Newey-West standard errors.

ferent socioeconomic groups of households' inflation expectations to explore their role in driving differing inflation expectations. In terms of forecasting power, the Bild indices perform better in predicting lower education and lower income households, while the opposite holds true for FAZ and SZ indices, which in general behave quite similar in comparison to the Bild indices. FAZ and SZ sentiments are additionally more suited to predict German HICP.

In a second step, I run local projections to estimate the impulse responses of the three newspaper indices to a monetary policy shock. Again, I can find heterogeneity in the responses: while the Bild sentiment reacts with an immediate and sizable decrease, suggesting a less negative tone towards the inflation environment, the response of FAZ and SZ indices is lagged by a few days and more muted. This could reflect the volatile and for mood-making aiming reporting style of the Bild also visible in the sentiment indices in Figure 4. The main policy implication I draw from this observation is that central banks should pay more attention to targeting tabloids in their communication strategy to take away their reasons for polarizing reporting as this is what some groups of households seem to resonate with. Another strategy could be to directly target those households that are most likely to be influenced by tabloids to prevent that they fill the void of information. This could be achieved by a higher frequency of easy understandable pieces of information shared, for example, on social media channels or television.

The last source of heterogeneity I explore is the varying responsiveness of different types of households to shocks in one of the three newspaper sentiment indices. In line

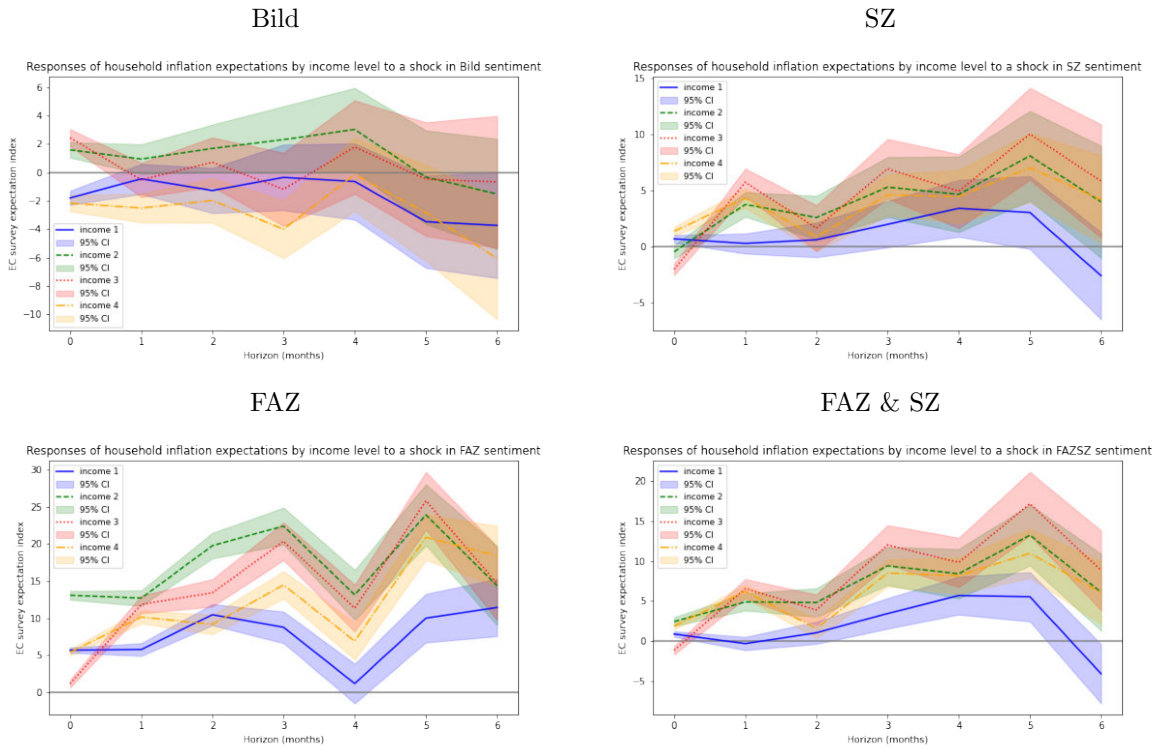


Figure 12: Cumulative impulse responses of EC survey expectations dis-aggregated across income quartiles to a shock in newspaper sentiment. Shaded areas represent 95% confidence bands constructed with Newey-West standard errors.

with my previous findings, expectations of households with a higher income and education level are most affected by a shock in FAZ and SZ sentiment. In the case of the Bild, this heterogeneity is less pronounced, but nonetheless, the strongest and significant effect happens for participants belonging to the second income quartile.

All in all, this paper shows that inflation sentiment indicators based on newspaper texts do not only carry useful real-time information for forecasting, but shocks to them also impact parts of the society differently. They therefore present themselves as a valuable tool for short-run predictions of heterogeneous household inflation expectations. My results also recommend central banks to pay more attention to public news reporting or to even steer their narratives on inflation by effective communication, as this can have immediate impact on the anchoring of inflation expectations. For the future, I expect such alternative inflation expectation measures to become more accurate and potent with advancements in text analytics with LLMs.

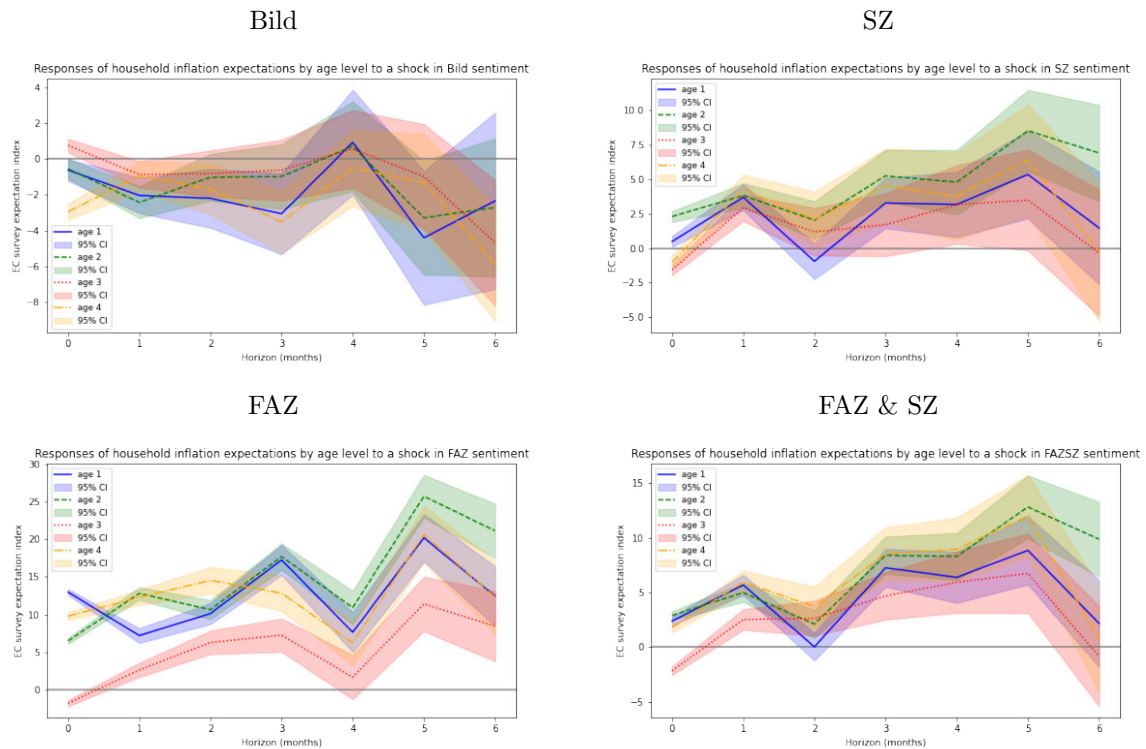


Figure 13: Cumulative impulse responses of EC survey expectations dis-aggregated across age groups to a shock in newspaper sentiment. Shaded areas represent 95% confidence bands constructed with Newey-West standard errors.

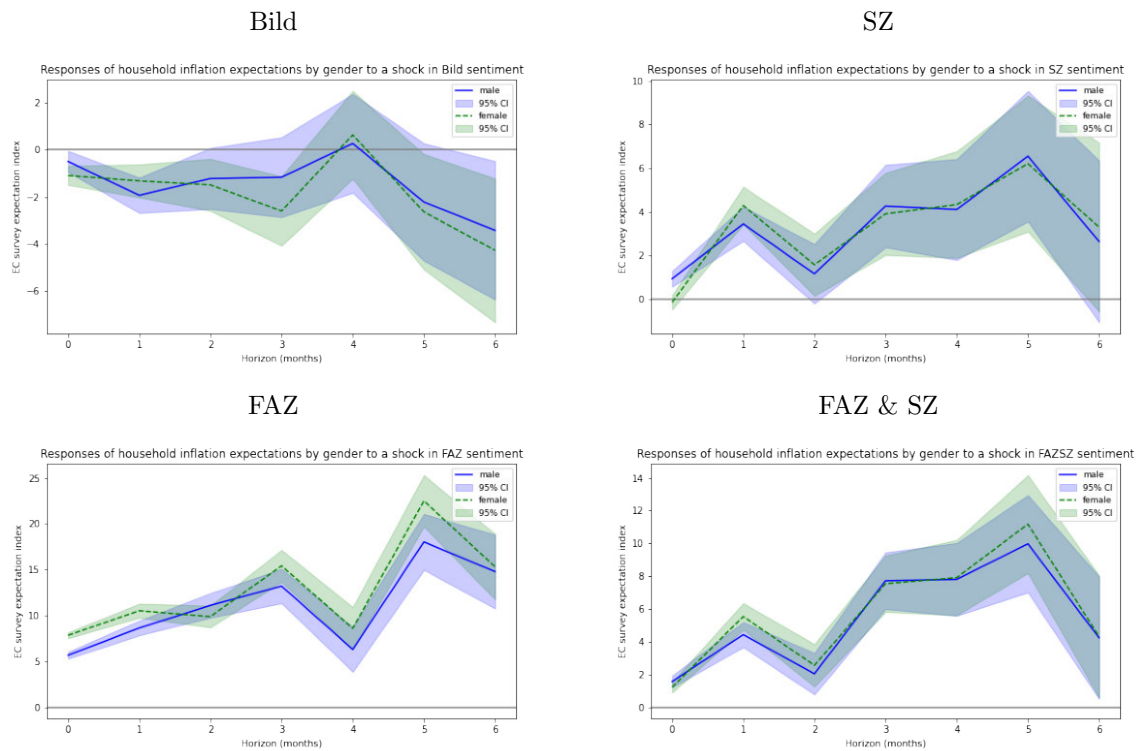


Figure 14: Cumulative impulse responses of EC survey expectations dis-aggregated across gender to a shock in newspaper sentiment. Shaded areas represent 95% confidence bands constructed with Newey-West standard errors.

References

- Altavilla, C., Brugnolini, L., Gürkaynak, R. S., Motto, R., and Ragusa, G. (2019). Measuring euro area monetary policy. *Journal of Monetary Economics*, 108:162–179.
- Angelico, C., Marcucci, J., Miccoli, M., and Quarta, F. (2022). Can we measure inflation expectations using twitter? *Journal of Econometrics*, 228(2):259–277.
- Bańkowska, K., Borlescu, A., Charalambakis, E., Dias da Silva, A., Di Laurea, D., Dossche, M., Georgarakos, D., Honkkila, J., Kennedy, N., Kenny, G., et al. (2021). Ecb consumer expectations survey: an overview and first evaluation. *ECB Occasional Paper*, (2021/287).
- Belal, M., She, J., and Wong, S. (2023). Leveraging chatgpt as text annotation tool for sentiment analysis. *arXiv e-prints*, pages arXiv–2306.
- Born, B., Dalal, H., Lamersdorf, N., and Steffen, S. (2023). Monetary policy in the age of social media: A twitter-based inflation analysis. *Available at SSRN*.
- Bybee, J. L. (2023). The ghost in the machine: Generating beliefs with large language models. *arXiv preprint arXiv:2305.02823*.
- Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *the Quarterly Journal of economics*, 118(1):269–298.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Ellingsen, J., Larsen, V. H., and Thorsrud, L. A. (2022). News media versus fred-md for macroeconomic forecasting. *Journal of Applied Econometrics*, 37(1):63–81.
- Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e. V. (IVW) (2024). Ivw digital. <https://www.ivw.de/>.
- Jarociński, M. and Karadi, P. (2020). Deconstructing monetary policy surprises - the role of information shocks. *American Economic Journal: Macroeconomics*, 12(2):1–43.
- Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review*, 95(1):161–182.
- Meeks, R. and Monti, F. (2023). Heterogeneous beliefs and the phillips curve. *Journal of Monetary Economics*, 139:41–54.
- Menz, J.-O. and Poppitz, P. (2013). Households’ disagreement on inflation expectations and socioeconomic media exposure in germany.

- Montiel Olea, J. L. and Plagborg-Møller, M. (2021). Local projection inference is simpler and more robust than you think. *Econometrica*, 89(4):1789–1823.
- Picault, M., Pinter, J., and Renault, T. (2022). Media sentiment on monetary policy: Determinants and relevance for inflation expectations. *Journal of International Money and Finance*, 124:102626.
- Shapiro, A. H., Sudhof, M., and Wilson, D. J. (2022). Measuring news sentiment. *Journal of econometrics*, 228(2):221–243.
- Wang, Z., Xie, Q., Ding, Z., Feng, Y., and Xia, R. (2023). Is chatgpt a good sentiment analyzer? a preliminary study. *arXiv preprint arXiv:2304.04339*.
- Weber, M., Candia, B., Ropele, T., Lluberas, R., Frache, S., Meyer, B. H., Kumar, S., Gorodnichenko, Y., Georgarakos, D., Coibion, O., et al. (2023). Tell me something i don't already know: Learning in low and high-inflation settings. Technical report, National Bureau of Economic Research.

Appendix

A Search queries

This section contains the detailed German versions of the search queries I use in the SZ and FAZ news archives, on LexisNexis and mediacloud to filter inflation related newspaper articles. On LexisNexis, Bild newspaper articles are available from 2010 to 2017. The Bild query uses a slightly different vocabulary such as *Zins-Hammer* (exaggeration for interest rate increase) or *Teuerschock* (i.e., a shock that increases prices by a lot) which is unique to this newspaper.

SZ:

Inflation* OR "niedrige Preise" OR "hohe Preise" OR Disinflation OR Deflation OR Geldpolitik OR Lohn-Preis-Spirale OR Zentralb* OR Stagflation OR Energiepreis* OR Geldsteuerung OR EZB OR Kaufkraftverlust OR Notenb* OR Leitzins OR Kostenanstieg OR Preissenkung* OR Zinswende OR Preisexplosion OR Preisanstieg OR Lebenshaltungskosten OR Teuerung OR Preisindex AND Deutschland NOT Verlag* NOT Preis-träger* NOT Preissieger* NOT *Film* NOT Oscar NOT Preisverleihung* NOT Preisgeld* NOT Billigung NOT Trainer* NOT Autor* NOT Fußball* NOT preisg* NOT Trump

FAZ:

(RESSORT: (Wirtschaft) OR RESSORT: (Politik) OR RESSORT: (Finanzmarkt) OR RESSORT: (Immobilienmarkt) NOT RESSORT: (Rhein-Main-Zeitung)) AND QUELLENANGABE: (Frankfurter Allgemeine Zeitung) AND (TEXT: (Inflation*) OR TEXT: (niedrige Preise) OR TEXT: (hohe Preise) OR TEXT:(Disinflation) OR TEXT: (Deflation) OR TEXT: (Geldpolitik) OR TEXT: (Lohn-Preis-Spirale) OR TEXT: (Zentralb*) OR TEXT: (Stagflation) OR TEXT: (Energiepreis*) OR TEXT: (Geldsteuerung) OR TEXT: (EZB) OR TEXT: (Kaufkraftverlust) OR TEXT: (Notenb*) OR TEXT: (Leitzins) OR TEXT: (Kostenanstieg) OR TEXT: (Preissenkung*) OR TEXT: (Zinswende) OR TEXT: (Preisexplosion) OR TEXT: (Preisanstieg) OR TEXT: (Lebenshaltungskosten) OR TEXT: (Teuerung) OR TEXT: (Preisindex) NOT (Verlag*)) AND ERSCHEINUNGSJAHR: [2010 TO 2023] AND LAND: (Deutschland) AND (QUELLENANGABE: (Frankfurter Allgemeine Zeitung) OR QUELLENANGABE: (FAZ.NET)) NOT SERIENTITEL: (Politische Bücher) NOT SERIENTITEL: (Wirtschaftsbücher)

LexisNexis:

Zins-Hammer or steigende Preise or fallende Preise or Preischaos or Bezinpreis* or Teuerschock or Inflation* or niedrige Preise or hohe Preise or Disinflation or Deflation or Geld-

politik or Lohn-Preis-Spirale or Zentralbank or Stagflation or Energiepreis* or Geldsteu-
erung or EZB or Kaufkraftverlust or Notenb* or Leitzins or Kostenanstieg or Preis-
senkung* or Zinswende or Preisexplosion or Preisanstieg or Lebenshaltungskosten or
Teuerung or Preisindex and not Verlag* and not Preisträger* and not Preissieger* and
not *Film* and not Oscar and not Preisverleihung* and not Preisgeld* and not Billigung
and not Trainer* and not Autor* and not Fußball* and not preisge* and not Trump

mediacloud:

(Zins-Hammer OR steigende Preise OR fallende Preise OR Preischaos OR Bezinpreis*
OR Teuerschock OR Inflation* OR niedrige Preise OR hohe Preise OR Disinflation
OR Deflation OR Geldpolitik OR Lohn-Preis-Spirale OR Zentralbank OR Stagflation
OR Energiepreis* OR Geldsteuerung OR EZB OR Kaufkraftverlust OR Notenban* OR
Leitzins OR Kostenanstieg OR Preissenkung* OR Zinswende OR Preisexplosion OR
Preisanstieg OR Lebenshaltungskosten OR Teuerung OR Preisindex) AND NOT (Verlag
OR Preisträger* OR Preissieger* OR Film OR Oscar OR Preisverleihung* OR Preisgeld*
OR Billigung OR Trainer* OR Autor* OR Fußball* OR preisge* OR Trump OR Tor*)

B ChatGPT’s interpretation of Bild expressions

ChatGPT is indeed able to understand the special language of the Bild newspaper and to capture the intended tone of its expressions. I demonstrate this by asking ChatGPT to give explanations and interpretations of the tone of the in Table 1 highlighted German terms, using the following prompt: *”Explain the term [...] in the context of the newspaper headline [...] and describe the tone this expression implies.”* Table B-1 presents the results of this exercise. In all cases, ChatGPT accurately captures the meaning and provides a suited interpretation of the German expressions.

Term	ChatGPT Explanation	Interpretation of tone
<i>Teuer-Schock</i>	The term "Teuer-Schock" is a German compound term that combines "Teuer," meaning expensive or costly, with "Schock," meaning shock. As a compound term, "Teuer-Schock" is often used in the context of economic or financial discussions to describe a situation where there is a sudden, unexpected, and significant increase in prices or costs, leading to a shock-like impact on individuals, businesses, or the economy as a whole.	The tone implied by "Teuer-Schock" is generally one of concern, surprise, and possibly criticism. It suggests a negative or adverse situation in which the cost or price surge is seen as problematic, potentially disruptive, or challenging for those affected. The term is used to convey the idea that the increase in prices is not just a regular or anticipated occurrence but rather an event that has a notable and adverse impact, similar to the impact of a shock. In news headlines or economic discussions, "Teuer-Schock" is often employed to draw attention to the severity and unexpectedness of the price increase, emphasizing its potential consequences and the need for attention or action.
<i>oberster Geld-Hüter</i>	The term "oberster Geld-Hüter" is a German expression that can be translated to "supreme money guardian" or "chief monetary guardian" in English. This term is often used to refer to a person who holds a key position of authority or responsibility in monetary or financial matters, such as a central bank governor, finance minister, or another high-ranking official overseeing economic policy.	The tone implied by "oberster Geld-Hüter" is one of authority, responsibility, and sometimes a degree of formality. It conveys the idea that the individual being referred to is at the highest level of decision-making and influence regarding financial and monetary policies. The term is often used in news articles, economic analyses, or discussions to highlight the significance and influence of the person in question, emphasizing their role in shaping and managing the country's economic policies. The use of such a term may carry a sense of expectations regarding the individual's ability to make important decisions and steer the country's economic course, reflecting a certain level of accountability and authority associated with the position.

Term	ChatGPT Explanation	Interpretation of tone
<i>Madame Inflation</i>	In the context of the euro zone, "Madame Inflation" could potentially refer to Christine Lagarde, the current President of the European Central Bank (ECB). Christine Lagarde holds a significant role in shaping and implementing monetary policies that influence economic conditions, including inflation rates, within the euro zone.	When using a term like "Madame Inflation" to refer to Christine Lagarde, it suggests a personification of responsibility for inflationary pressures or challenges within the euro zone. The term implies a critical or at least skeptical stance toward her policies or actions, holding her accountable for the economic conditions that may lead to inflation. The use of "Madame" adds a formal and somewhat sardonic tone, suggesting a level of authority and influence associated with Christine Lagarde's position. Overall, it implies a critical perspective on the individual in charge of monetary policies in the euro zone, particularly in the context of concerns or shocks related to inflation.
<i>Luxus-Lagarde</i>	In the context of the newspaper headline "Madame Inflation; Luxus-Lagarde macht Sparer und Rentner arm," the term "Luxus-Lagarde" is a metaphorical expression that combines "Luxus" (luxury) with the surname "Lagarde," referring to Christine Lagarde, the President of the European Central Bank (ECB).	The term "Luxus-Lagarde" suggests a critical tone, implying that Christine Lagarde is associated with a lifestyle of luxury or extravagance. By combining her name with "Luxus," it conveys the idea that Lagarde is perceived as leading a lifestyle that may be considered opulent or out of touch with the financial challenges faced by the general population, specifically savers and retirees mentioned in the headline. The tone of "Luxus-Lagarde" in this context is likely critical, implying a disapproval of Lagarde's perceived lifestyle in connection with the economic challenges faced by savers and retirees. It may be used as a form of criticism, suggesting a disconnect between Lagarde's perceived luxurious lifestyle and the economic struggles of certain groups in society. This kind of expression is often employed in headlines or articles to convey a strong opinion or perspective on a public figure's actions or policies.

Term	ChatGPT Explanation	Interpretation of tone
<i>Zins-Hammer</i>	In the context of the newspaper headline "Zins-Hammer! EZB erhöht auf 1,25 Prozent" (Interest Rate Hammer! ECB Increases to 1.25 Percent), the term "Zins-Hammer" is a metaphorical expression. It combines "Zins" (interest rate) with "Hammer," suggesting a forceful and impactful event related to interest rates.	The term "Zins-Hammer" implies a tone of surprise, emphasis, and potentially a sense of urgency or significance. It suggests that the increase in interest rates is not just a routine adjustment but a substantial and noteworthy event that is likely to have important consequences or effects. The use of "Hammer" adds a dramatic and attention-grabbing element to the headline, indicating that the interest rate change is being portrayed as a powerful or forceful action. It might also imply that the change in interest rates could have considerable impacts on various stakeholders, such as borrowers, savers, and the overall economy.

Table B-1: Explanations and interpretations of Bild-typical expressions shown in Table 1 generated by ChatGPT in response to the following prompt: "Explain the term [...] in the context of the newspaper headline [...] and describe the tone this expression implies."

C Further correlation heat maps

This section presents the remaining sample correlation heat maps of the SZ expectation indices and the EC and ECB surveys (Figure C-1), as well as different sample splits for the EC survey data (2014M01-2020M03 and 2020M04-2023M08) and the three newspapers in Figure C-2. Overall the correlation in the latter sample split is higher. The pattern of heterogeneity between different groups of participants remains consistent for the Bild indices, while it somewhat reverses for the FAZ and SZ indices.

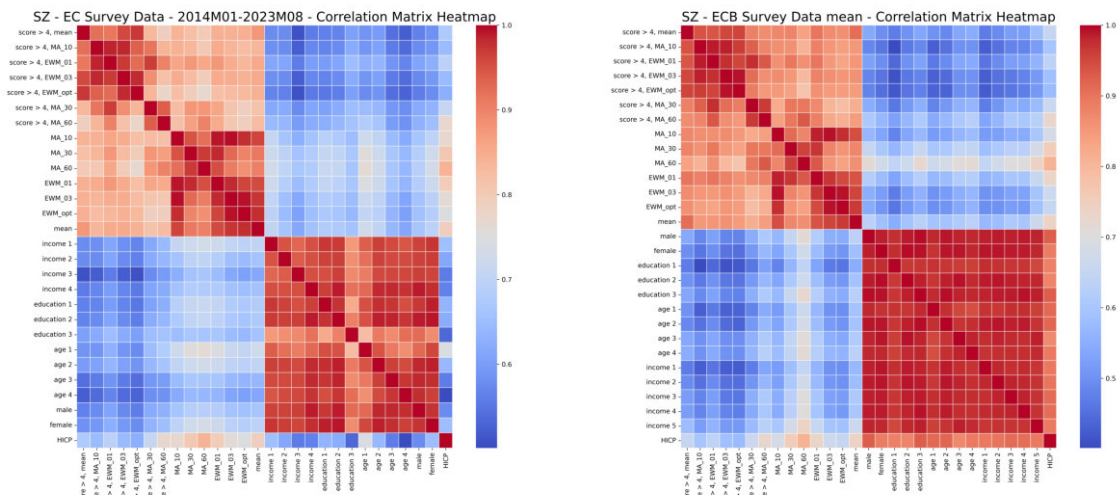


Figure C-1: Sample correlations between SZ inflation indices and EC data (left) and ECB survey data (right) based on the whole sample period; 2014M01-2023M08 for EC data and 2020M04-2023M08 for ECB data. Red indicates high and blue low correlations.

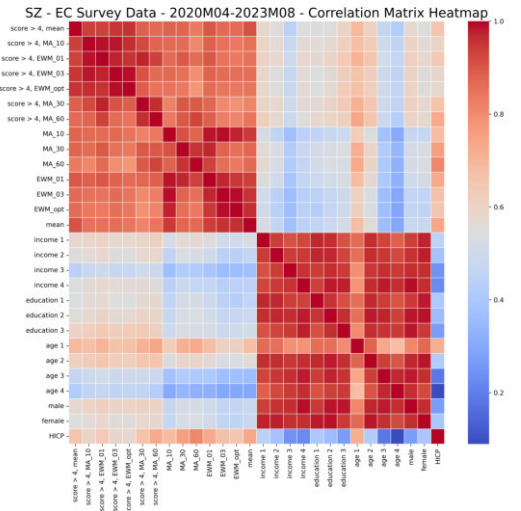
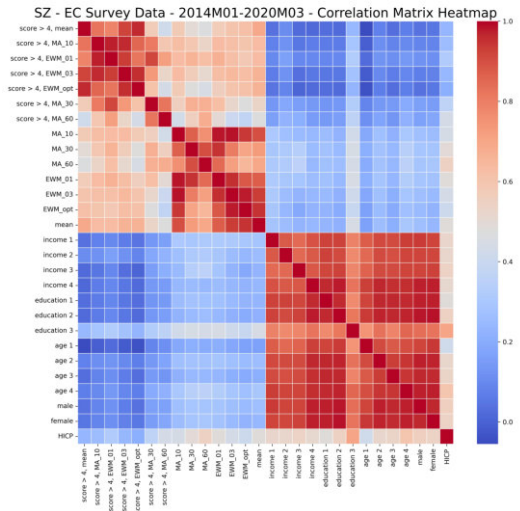
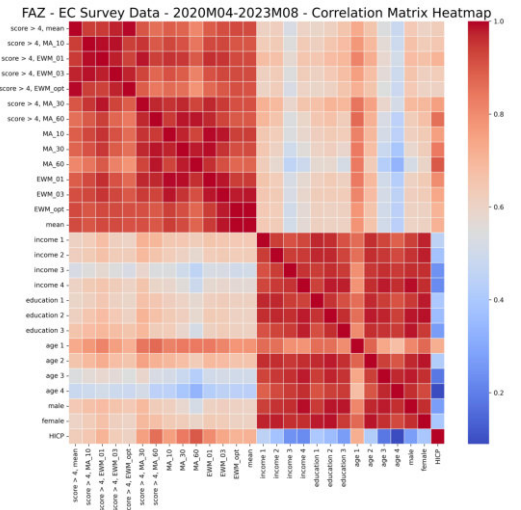
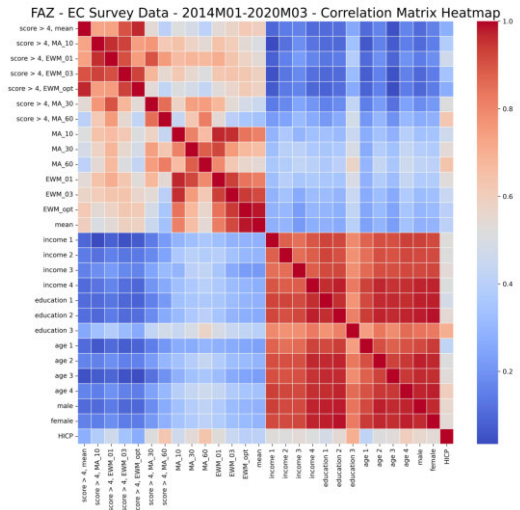
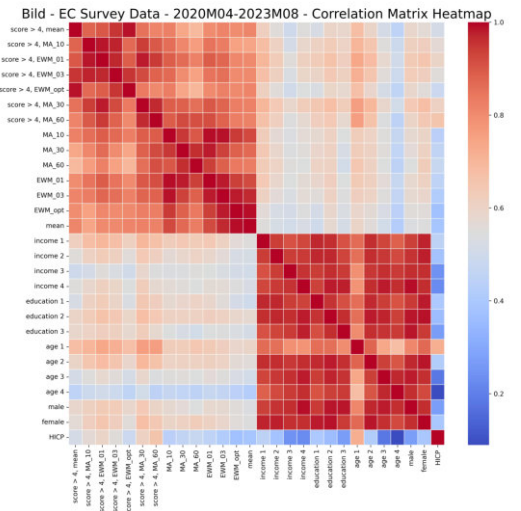
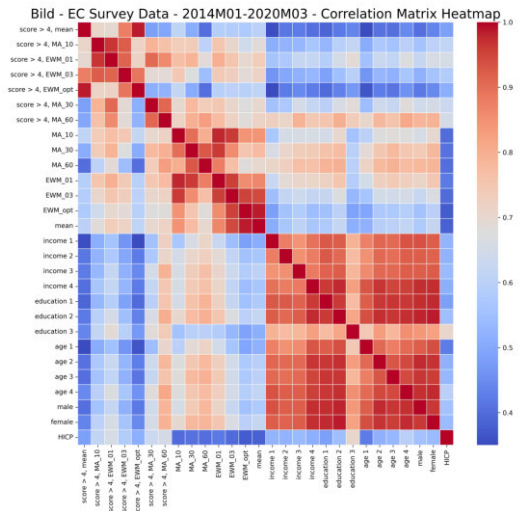


Figure C-2: Sample correlations between Bild, FAZ and SZ (top to bottom) inflation indices and EC data on the 2014M01-2020M03 (left) and the 2020M04-2023M08 sample period (right). Red indicates high and blue low correlations.